

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

M nucleic - protein search, using frame_plus_n2p model

un on: October 9, 2003, 12:12:11 ; Search time 78 Seconds
(without alignments)
2779.028 Million cell updates/sec

title: US-09-518-842-1

effect score: 774

sequence: 1 ATGGCCAGCTGTCCGGT.....CAGTTAAATATGATACATAG 420

coring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

searched: 830525 seqs, 258052604 residues

total number of hits satisfying chosen parameters: 1661050

inimum DB seq length: 0

aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

command line parameters:

MODEL=frame+n2p,model -DEV=xlp
QD=/cgn2/1/USPTO.spool/US09518842/runat_09102003_111036/app_query.fasta_1.583
DB=SPREMBL_23 -QFMT=fastan -SUFFIX=rspt -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
DOCALLIGN=200 -THR SCORE=ptc -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
DOFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
JSER=US09518842 @CGN 1 172 @runat_09102003_111036_13065 -NCPU=6 -ICPU=3
VO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT DSPBLOCK=100 -LONGLOG
DEV_TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

atabase :

- 1: sp archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

result	Query	Score	Match	Length	ID	Description
1	154.5	20.0	199	6	Q9N0T9	Q9N0T9 camelus dro

ALIGNMENTS

RESULT 1

ID	Q9N0T9	PRELIMINARY;	PRT;	199 AA.
AC	Q9N0T9;			
DT	01-OCT-2000 (TRENBLrel. 15, Created)			
DT	01-OCT-2000 (TRENBLrel. 15, Last sequence update)			
DT	01-MAR-2003 (TRENBLrel. 23, Last annotation update)			
DE	Preprorelaxin.			
OS	Camelus dromedarius (Dromedary) (Arabian camel).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.			
OX	NCBI_TaxID=9838;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20191540; PubMed=10727251;			
RA	Hombach-Klonisch S., Abd-Elnaeem M., Skidmore J.A., Leiser R.,			
RA	Fischer B., Klonisch T.;			
RT	"Ruminant relaxin in the pregnant one-humped camel.";			
RL	Biol. Reprod. 62:839-846(2000).			
CC	-!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).			
CC	-!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.			
DR	EMBL; AF254739; AAF67741.1; -			
DR	InterPro; IPR004825; Ins/IGF/relax.			
DR	Pfam; PF00049; Insulin; 1.			
DR	SMART; SM00078; IIGF; 1.			

Q9MYK8 felis silve
Q9K46 galago cras
Q22328 caenorhabdi
Q8BT12 mus musculu
Q8ZQC3 salmonella
Q8Z802 salmonella
Q8TC33 homo sapien
Q9NUP6 homo sapien
Q9RLW4 klebsiella
Q8CIS1 mus musculu
Q8CIDI mus musculu
Q9HZB9 pseudomonas
Q9BE45 bos taurus
Q64293 streptococc
Q25340 leishmania
Q8CG10 mus musculu
Q9DBY0 mus musculu
Q8KOP9 mus musculu
Q961W9 drosophila
Q8CDN3 mus musculu
Q36168 streptococc
Q8P500 xanthomonas
Q8ZB58 yerinia pe
Q8L712 arabidopsis
Q8DI0 yerinia pe
Q9FHH9 arabidopsis
Q9B2B5 homo sapien
Q8RUD8 oryza sativ
Q8CB10 mus musculu
Q8CIP5 mus musculu
Q8MKE1 macropus eu
Q8EYX2 leptospira
P91457 caenorhabdi
Q9D718 mus musculu
Q9ATG6 polytomella
Q19057 pongo pygma
Q89241 western equ
Q88791 western equ
Q8NDB1 homo sapien
Q82202 arabidopsis
Q8IV53 homo sapien
Q70148 rattus norv
Q8NC88 homo sapien
Q96JM3 homo sapien

C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
C Mammalia; Eutheria; Primates; Strepsirhini; Galagonidae; Otollemur.

X NCBI_TaxID=9463;

N [1]

N P SEQUENCE FROM N.A.

C TISSUE=Placenta;

A Klonisch T., Froehlich C., Tetens F., Fischer B., Hombach-Klonisch S.;

T "Molecular remodeling of members of the relaxin family during primate

L evolution.";

T Mol. Biol. Evol. 0:0-0(2001).

C -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).

C -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.

R EMBL; AF317625; AAC42318.1; -

R HSSP; P04090; 6RLX.

R InterPro; IPR004825; Ins/IGF/relax.

R Pfam; PF00049; Insulin; 1.

R SMART; SM00078; IIGF; 1.

R PROSITE; PS00262; INSULIN; 1.

Q SEQUENCE 188 AA; 21573 MW; 77551629B82B9A66 CRC64;

Alignment Scores:

red. No.: 0.00772 Length: 188

core: 103.00 Matches: 48

Percent Similarity: 34.50% Conservative: 21

Best Local Similarity: 24.00% Mismatches: 57

Query Match: 13.31% Indels: 74

B: 6 Gaps: 9

S-09-518-842-1 (1-420) x Q9GK46 (1-188)

Y 1 ATGCCAGCTGTTCGGTCTATCTGCCAGCAATCTGGCTGCTGAGCCAACTCCTT 60

b 1 MetProArgLeuPhePheHisLeuLeuGlyValTrpLeuLeuThrGlnIleSer 20

Y 61 AGA-----GAAAGCTAGCAGCAGAGCTGAGGGATGTGTCCCA 102

b 21 ArgAlaLysMetAspLysGlyGluAsnLeuAsnGlnIlePheAlaCysGlyArg 40

Y 103 TTGGAAACACTTGTCTCATATTCCTCCATGCTCCGAGAGACATTCACCCACCCCA 162

b 41 LeuIleArgIleTrpValGluValCys----- 49

Y 163 GGAGGTGGCTGTGGAATCTGACGCTCCAAA----- 195

D 50 ---GlySerThrGlyPheArgGlyArgAlaLysAsnGlnThrGluHisGlnProGlySer 68

Y 196 -----CAATGGTGTCACTCC-----NACAACAGATGCAGACGCTTAGGT 240

b 69 GluProPheSerGluIleValProSerSerPheIleAsnLysAspAlaGluThrIleAsn 88

Y 241 ACACATCAGAAATTCATTCCTAATTTG-----CCACTGCTGAAGG----- 300

b 89 MetSerGluPheIleAlaAsnLeuProGlnLysGlnLysThrGlnSerGluMet 108

Y 268 -----TCACAGAGCTCAGAAA----- 300

b 109 AsnLeuProSerProGluLeuGlnGlnTyProProThrLeuLysGlySerAspIleSer 128

Y 301 -----CAG 303

b 129 PheGluGluValLysAsnAsnIleHisAsnGluGlnGlyGluAlaGluAspAsnSerHis 148

Y 304 CCATCATTTGAAGAAATAATACTTCCCGAAAAAGAGAGTGGAGCTCACAGATT--- 360

b 149 SerGluLeuGlnAsnLeuGlyLeuAspThrHisSerArgLysLysArgGluArgTyrMet 168

Y 361 GATCCATTC-----TGTTGGAGTAATTTGTGACGATGGAACTTCAGTTAATATGT 414

b 169 SerProLeuGlnLysCysCysArgIleGlyCysThrLysArgSerLeuAlaArgPheCys 188

RESULT 4

22328

D Q22328 PRELIMINARY; PRT; 575 AA.

AC Q22328;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Hypothetical protein T07H6.5.

GN T07H6.5.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RX MEDLINE=99069613; PubMed=9851916;

RA Waterston R.;

RT "Genome sequence of the nematode C. elegans: a platform for

RL investigating biology. The C. elegans Sequencing Consortium.";

RL Science 282:2012-2018(1998).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RA Geisler C.;

RT "The sequence of C. elegans cosmid T07H6.5";

RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RA Waterston R.;

RL EMBL; U53344; AAA96225.2; -

DR HSSP; P10998; 1VVD.

DR WormPep; T07H6.5; CE31364.

DR InterPro; IPR000436; Sushi_SCR_CCP.

DR Pfam; PF00084; sushi; 8.

DR SMART; SM00032; CCP; 8.

KW Hypothetical protein.

SQ SEQUENCE 575 AA; 63326 MW; B6820F6C692AA646 CRC64;

Alignment Scores:

Pred. No.: 0.0171 Length: 575

Score: 101.00 Matches: 43

Percent Similarity: 38.56% Conservative: 16

Best Local Similarity: 28.10% Mismatches: 38

Query Match: 13.05% Indels: 56

DB: 5 Gaps: 9

US-09-518-842-1 (1-420) x Q22328 (1-575)

QY 94 GGTCCCGGATTGGAAAACACTTCTCTCATATTGCCCCCATGCTGAGAGACATTCAAC 153

Db 429 GlyProMet---LysCysLeuProSerTrpCysGluHisProSerLysThrTyrGly 447

QY 154 ACCACCCGAGGGGTGGCTGCTG----- 177

Db 448 ThrLeuProGlyGlyGlnIleLeuLeuGluGlyIleLeuGlyAlaTyrGluPheGlnSer 467

QY 178 -----GAATCTGGAGCTGCCAACAAAGAAATGGTGTCAACCTCCCAACAAA 222

Db 468 TyrIleGlnLysValGluGluGlyArg-----AlaIleSerPheGlnCysGlyLys 484

QY 223 GATGACAAAGCCTTAGGT-----ACGACATCAGAAATTCATTCCTTAAT 264

Db 485 GlyAsnTyrLeuIleGlyProProLysAlaThrCysValAsnGlyGluTrpMetProLys 504

QY 265 TTGTACACGAGCTGAAGAACCACTGCTGAGGGCAGGCATCATTTGAAG---AAATA 321

Db 505 ValSerPro-----LysCysValSerGlnThrHisProMetIleGluGlyLysIle 521

QY 322 ATACTTTCCTCCCAAAAGAGAGT-----GGAGCT----- 351

Db 522 LeuTrpAspArgLysLysArgSerLeuProGlyArgAlaValArgGluTyrValAspAsp 541

QY 352 -----CACAGATTGTGATTCATTCTGTTGTAAGTAATT----- 384


```
Q9NUP6
ID Q9NUP6 PRELIMINARY; PRT; 754 AA.
AC Q9NUP6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein FL011220.
OS Homo sapiens (Human).
JC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
DC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA Ninomiya K., Iwayanagi T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
KW Hypothetical protein.
SQ SEQUENCE 754 AA; 85004 MW; 5966323DAE4F5B67 CRC64;

Alignment Scores:
Pred. No.: 0.494 Length: 754
Score: 88.50 Matches: 39
Percent Similarity: 44.6% Conservative: 19
Best Local Similarity: 30.0% Mismatches: 50
Query Match: 11.4% Indels: 22
Gaps: 7

S-09-518-842-1 (1-420) x Q9NUP6 (1-754)
Y 9 CTTGTCGGCTCTATCTGCCAGCAATCTGGCTGCTGCTGAGCCACTCTTACAGAAAG 68
b 99 ProProProSerLeuAlaProAlaGlyProAlaValAlaAlaProLeuPro-----Ala 116
Y 69 CTTAGCAGCAGCTGAGGGGATGCTGCCCGATTGCGAAACACTTGTGTCATATTG 128
b 117 ProSerThrSerAlaLeuPheThrPheSerProLeu-----ThrValSerAlaAlaGly 134
Y 129 CCC-----CATGCTGAGAACACATTCACACACCCCGAGGGGTGGCTGCT 176
b 135 ProLysHisLysGlyHisLysGluArgHisLysHisLysHisLysHisLysArg-Gly----- 151
Y 177 GGAATCTGACGCTCCCAAGAAATGCTGCACTCCCAACCAACAAAGATGACAGCCTT 236
b 152 -ProAspGlyAspProSerSerCysGlyThrAspLeuLysHisLysAspLysGlnGluAs 171
Y 237 AGGTACGACATCAGATTCATTCCTAATTG-----TCACAGAGCTGAAGAACCACTGTC 293
b 171 nGlyGluArgThrGlyGlyValProLeuIleLysAlaProLysArgGluThrProAspGly 191
Y 294 TGAAGGGCAGCCATCA-----TTGAAGAAATAATACTTTCCCGCAA 335
b 191 uAsnGlyLysThrGlnArgAlaAspPheValLeuLysLysLysLysLysLysLysLys 208
Y 336 AAAGAGAGCTGGAGCTCAGCATTTGAT 363
b 208 sLysLysLysLysLysLysHisArgGluAsp 217

RESULT 10
Q9RLM4 PRELIMINARY; PRT; 1218 AA.
D Q9RLM4;
C Q9RLM4;
T 01-MAY-2000 (TrEMBLrel. 13, Created)
T 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
T 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
E Virtulence protein S.

GN KVS.
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Klebsiella.
OX NCBI_TaxID=573;
RN [1]
RP SEQUENCE FROM N.A.
RC Yang S.L., Chang H.Y., Peng H.L.;
RT "Identification and characterization of KvgASOR, a two component
RT signal transduction system, in Klebsiella pneumoniae.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO OTHER PROKARYOTIC SENSORY TRANSDUCTION HISTIDINE
CC KINASES.
CC EMBL; AJ250891; CAB61240.1; -.
CC HSSP; P06143; IAB6.
DR InterPro; IPR003594; ATPbind ATPase.
DR InterPro; IPR004358; Bact_sens_pr_C.
DR InterPro; IPR003661; His_KinA.
DR InterPro; IPR005467; His_kinase.
DR InterPro; IPR000014; PAS_domain.
DR InterPro; IPR001789; Response_reg.
DR InterPro; IPR001311; SBP_glu_receptor.
DR InterPro; IPR001638; SBP_bac_3.
DR Pfam; PF02518; HATPase_c; 1.
DR Pfam; PF00512; HisKA; 1.
DR Pfam; PF00072; response_reg; 1.
DR Pfam; PF00497; SBP_bac_3; 1.
DR PRINTS; PR00344; BCTRLSENSOR.
DR ProDom; PD000039; Response_reg; 1.
DR SMART; SM00387; HATPase_c; 1.
DR SMART; SM00388; HisKA; 1.
DR SMART; SM00091; PAS; 1.
DR SMART; SM00062; PBPD; 2.
DR SMART; SM00448; REC; 1.
DR PROSITE; PS0109; HIS_KIN; 1.
DR PROSITE; PS0110; RESPONSE_REGULATORY; 1.
KW Kinase; Phosphorylation; Sensory transduction; Transferase.
SQ SEQUENCE 1218 AA; 136317 MW; 24F5889F2B9F02C0 CRC64;

Alignment Scores:
Pred. No.: 0.721 Length: 1218
Score: 87.50 Matches: 30
Percent Similarity: 44.9% Conservative: 19
Best Local Similarity: 27.5% Mismatches: 47
Query Match: 11.3% Indels: 13
Gaps: 3

US-09-518-842-1 (1-420) x Q9RLM4 (1-1218)
QY 7 AGCTGTGTC-----CGGTCCTATCTGCCAGCAATCTGGCTGCTGCTGAGCAACTCCTT 60
b 620 SerLeuPheAspArgSerProLeuValHisLePheProValIleGlnGlnMet 639
QY 61 AGAGAAAGCTTAGCAGCAGAGCTG-----AGGGGATGT 93
b 640 GlnGlnGlyLeuAlaProGluThrValAlaSerHisGlnPheThrLeuAsnAsnGlyAla 659
QY 94 GGTCCCGGATTTGGAAACACTTGTGTCATATTGCCCTGCTGAGACACATTCACC 153
b 660 GluAspArgThrIleLeuHisTrpMetThrLeuCysProLeuProValProLeuPro 679
QY 154 ACCACCCCGAGGGGTGG-----CTGCTGGAATCTGGACGCTCCCAAGAAATGGTGCA 207
b 680 ValLeuIleCysGlyTrpGlnAspIleThrGluSerArgLysLeuMetGluAlaLeuGln 699
QY 208 ACCTCCACACAAAGATGGACAGCCTTAGTACGACATCAGATTCATTCCTAAATTG 267
b 700 ValGluLysAspLysAlaIleGluAlaSerArgAlaLysSerArgPheLeuAlaArgMet 719
QY 268 TCACGAGAGCTGAAGAACCACTGTCT 294
b 720 SerHisGluIleArgThrProValSer 728
```



```

bb      1111 PheTrpSerLysIleTrpAsnThrThr----- 1120
yy      150 CACCACACCCAGGAGGGTGGCTGCTGGAAATCTGGACGTCCTCCCAAGAAATGGTGTCAAC 209
b      1121 -----AsnValAlaGlySerValTrpAsnGlyLeuGlyAsnGlyValAsp 1135

```

search completed: October 9, 2003, 12:18:38
 ob time : 84 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

DM nucleic - protein search, using frame_plus_n2p model

Run on: October 9, 2003, 12:13:06 ; Search time 21 Seconds
(without alignments)
1692.434 Million cell updates/sec

File: US-09-518-842-1

Perfect score: 7/4

Sequence: 1 ATGCCACGCTGTTCCGGTC.....CAGTTAAATTATGACATAG 420

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 657434

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

MODEL=frame+ n2p.model -DEV=xlp
Q=/cgn2_1/USPTO.spool/US09518842/runat_09102003_111038_13105/app_query.fasta_1.583
DB=Issued Patents AA -QPWT=fastran -SUFFIX=rai -MINMATCH=0.1 -LOPCL=0
LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=numar40.cdi
LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
MODE=LOCAL -OUTPMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
USER=US09518842 @CGN 1 1 38 @runat_09102003_111038_13105 -NCPU=6 -ICPU=3
NO MAP -LARGESQRY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:
1: /cgn2_6/ptodata/1/iaa/5A COMB.pcp.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pcp.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pcp.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pcp.*
5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pcp.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

result	No.	Score	Query Match	Length	ID	Description
1	739	95.5	139	3	US-08-950-720A-15	Sequence 15, Appl
2	739	95.5	139	3	US-08-991-890-2	Sequence 2, Appli
3	733	94.7	139	4	US-09-599-564A-2	Sequence 2, Appli
4	733	94.7	174	3	US-09-174-465D-2	Sequence 4, Appli
5	620	80.1	124	3	US-08-991-890-5	Sequence 5, Appli
6	519	80.0	159	3	US-08-991-890-4	Sequence 4, Appli
7	255	32.9	51	3	US-09-174-465D-10	Sequence 10, Appl
8	255	32.9	51	4	US-09-599-564A-10	Sequence 10, Appl
9	228	29.5	41	3	US-09-174-465D-8	Sequence 8, Appli
10	228	29.5	41	4	US-09-599-564A-8	Sequence 8, Appli
11	228	29.5	41	4	US-09-201-227A-27	Sequence 27, Appl
12	171	22.1	30	3	US-09-174-465D-12	Sequence 12, Appl

13	171	22.1	30	4	US-09-599-564A-12	Sequence 12, Appl
14	147	19.0	25	4	US-09-201-227A-28	Sequence 28, Appl
15	123.5	16.0	185	3	US-08-950-720A-12	Sequence 12, Appl
16	119.5	15.4	185	3	US-08-950-720A-13	Sequence 13, Appl
17	113.5	14.7	150	1	US-08-443-568B-12	Sequence 12, Appl
18	113.5	14.7	150	5	PCT-US94-06997-12	Sequence 12, Appl
19	113.5	14.7	162	6	5464756-18	Patent No. 5464756
20	101.5	13.1	164	6	5464756-20	Patent No. 5464756
21	85	11.0	17	3	US-09-174-465D-6	Sequence 6, Appli
22	85	11.0	17	4	US-09-599-564A-6	Sequence 6, Appli
23	83	10.7	324	4	US-09-252-991A-28877	Sequence 28877, A
24	80	10.3	322	6	5212074-1	Patent No. 5212074
25	75	10.1	994	3	US-08-699-103B-3	Sequence 3, Appli
26	75	10.1	994	4	US-09-229-059-3	Sequence 3, Appli
27	75	10.1	994	4	US-09-628-133-3	Sequence 3, Appli
28	75	10.1	995	4	US-08-747-582-15	Sequence 15, Appl
29	75	10.1	1711	2	US-08-342-930-2	Sequence 2, Appli
30	74	9.6	206	4	US-09-252-991A-34085	Sequence 24085, A
31	73	9.4	458	4	US-08-618-485B-1	Sequence 1, Appli
32	73	9.4	458	6	5177002-1	Patent No. 5177002
33	71.5	9.2	778	4	US-09-556-877-193	Sequence 193, App
34	71.5	9.2	778	4	US-09-620-412C-193	Sequence 193, App
35	71.5	9.2	778	4	US-09-598-419-193	Sequence 193, App
36	71.5	9.2	1530	4	US-09-556-877-178	Sequence 178, App
37	71.5	9.2	1530	4	US-09-620-412C-178	Sequence 178, App
38	71.5	9.2	1530	4	US-09-598-419-178	Sequence 178, App
39	70	9.0	344	1	US-08-843-993-3	Sequence 3, Appli
40	70	9.0	344	3	US-09-059-520A-3	Sequence 3, Appli
41	70	9.0	344	3	US-09-334-275-3	Sequence 3, Appli
42	69.5	9.0	754	4	US-09-252-991A-33133	Sequence 33133, A
43	69	8.9	239	5	PCT-US93-01652-1	Sequence 1, Appli
44	69	8.9	458	6	5177002-2	Patent No. 5177002
45	69	8.9	474	1	US-08-222-619-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-08-950-720A-15
; Sequence 15, Application US/08950720A
; Patent No. 6046028
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell C.
; APPLICANT: Lofton-Day, Catherine E.
; APPLICANT: Lok, Si
; APPLICANT: Jaspers, Stephen R.
; TITLE OF INVENTION: INSULIN HOMOLOG
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/950,720A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sawislak, Deborah A
; REGISTRATION NUMBER: 37,438
; REFERENCE/DOCKET NUMBER: 96-09
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6672

```

; TELEFAX: 206-442-6678
; TELEX:
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 139 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: NO. 6046028e
; US-08-950-720A-15

```

```

Alignment Scores:
Pred. No.: 4,05e-84 Length: 139
Score: 739.00 Matches: 139
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 95.48% Indels: 0
DB: 3 Gaps: 0

```

```

US-09-518-842-1 (1-420) x US-08-950-720A-15 (1-139)

```

```

QY 1 ATGCCAGCTGTCGGTCTCTATCGCAGCAATCTGCTGCTGCTGAGCCAACTCCTT 60
DB 1 MetAlaSerLeuPheArgSerTyrLeuProAlaIleTrpLeuLeuSerGlnLeuLeu 20
QY 61 AGAGAAAGCCTAGCAGCAGCTGAGGGATGTTGGTCCCGGATTTGGAAACACTTGCTG 120
DB 21 ArgGluSerLeuAlaIleLeuArgGlyCysGlyProArgPheGlyLysHisLeuLeu 40
QY 121 TCATATTGCCCATGCTGAGAGACATTCACCAACCCAGGAGGTGGTCTGGAA 180
DB 41 SerTyrCysProMetProGluLysThrPheThrThrProGlyGlyTrpLeuLeuGlu 60
QY 181 TCTGAGCGTCCCAAGAAATGTGTCAACCTCCACACAAAGATGGACAGCCTTAGT 240
DB 61 SerGlyArgProLysGluMetValSerThrSerAsnLysAspGlyGlnAlaLeuGly 80
QY 241 ACACATCAGATTCTTCTTAATTTGTCCAGAGCTGAACAAACCACTGTCTGAAGG 300
DB 81 ThrThrSerGluPheIleProAsnLeuSerProGluLeuLysProLeuSerGluGly 100
QY 301 CAGCCATCATTTGAAGAAATAATCTTTCCCGCAAAAAGAGAGTGGACGTACAGATT 360
DB 101 GlnProSerLeuLysLysIleLeuSerArgLysLysArgSerGlyArgHisArgPhe 120
QY 361 GATCCATTCTGTGTCAAGTAATTTGTGACGATGGAACTTCAGTTAAATTTATGTACA 417
DB 121 AspProPheCysCysGluValIleCysAspAspGlyThrSerValLysLeuCysThr 139

```

RESULT 2

```

JS-08-991-890-2
Sequence 2, Application US/08991890
Patent No. 6114307

```

GENERAL INFORMATION:

```

; APPLICANT: Jaspers, Stephen R.
; APPLICANT: Sprugel, Katherine H.
; APPLICANT: Ren, Hong Ping
; APPLICANT: Humes, Jacqueline M.
; APPLICANT: Hoffman, Ross C.
; APPLICANT: Conklin, Darrell C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: STIMULATING PANCREATIC ISLET CELL REGENERATION
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible

```

```

; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/991,890
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/033,003
; FILING DATE: December 16, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sawislak, Deborah A
; REGISTRATION NUMBER: 37,438
; REFERENCE/DOCKET NUMBER: 96-41
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6672
; TELEFAX: 206-442-6678
; TELEX:

```

```

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 139 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-991-890-2

```

```

Alignment Scores:
Pred. No.: 4,05e-84 Length: 139
Score: 739.00 Matches: 139
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 95.48% Indels: 0
DB: 3 Gaps: 0

```

```

US-09-518-842-1 (1-420) x US-08-991-890-2 (1-139)

```

```

QY 1 ATGCCAGCTGTCGGTCTCTATCGCAGCAATCTGCTGCTGCTGAGCCAACTCCTT 60
DB 1 MetAlaSerLeuPheArgSerTyrLeuProAlaIleTrpLeuLeuSerGlnLeuLeu 20
QY 61 AGAGAAAGCCTAGCAGCAGCTGAGGGATGTTGGTCCCGGATTTGGAAACACTTGCTG 120
DB 21 ArgGluSerLeuAlaIleLeuArgGlyCysGlyProArgPheGlyLysHisLeuLeu 40
QY 121 TCATATTGCCCATGCTGAGAGACATTCACCAACCCAGGAGGTGGTCTGGAA 180
DB 41 SerTyrCysProMetProGluLysThrPheThrThrProGlyGlyTrpLeuLeuGlu 60
QY 181 TCTGAGCGTCCCAAGAAATAATCTTTCCCGCAAAAAGAGAGTGGACGTACAGATT 240
DB 61 SerGlyArgProLysGluMetValSerThrSerAsnLysAspGlyGlnAlaLeuGly 80
QY 241 ACACATCAGATTCTTCTTAATTTGTCCAGAGCTGAACAAACCACTGTCTGAAGG 300
DB 81 ThrThrSerGluPheIleProAsnLeuSerProGluLeuLysProLeuSerGluGly 100
QY 301 CAGCCATCATTTGAAGAAATAATCTTTCCCGCAAAAAGAGAGTGGACGTACAGATT 360
DB 101 GlnProSerLeuLysLysIleLeuSerArgLysLysArgSerGlyArgHisArgPhe 120
QY 361 GATCCATTCTGTGTCAAGTAATTTGTGACGATGGAACTTCAGTTAAATTTATGTACA 417
DB 121 AspProPheCysCysGluValIleCysAspAspGlyThrSerValLysLeuCysThr 139

```

RESULT 3

```

US-09-599-564A-2
Sequence 2, Application US/09599564A
Patent No. 6362318
; GENERAL INFORMATION:
; APPLICANT: KOMAN Ahment
; APPLICANT: CHASSIN, Dorine
; APPLICANT: BELLET, Dominique

```

TITLE OF INVENTION: NEW PROTEIN CALLED EPIL/PLACENTIN, PROCESS FOR THE
TITLE OF INVENTION: PREPARATION OF THIS PROTEIN AND PHARMACEUTICAL
TITLE OF INVENTION: COMPOSITION CONTAINING SUCH, DNA CODING FOR SAID
FILE REFERENCE: 017753-127

CURRENT APPLICATION NUMBER: US/09/599,564A

PRIOR FILING DATE: 2000-06-23

PRIOR FILING DATE: 09/174,465

PRIOR FILING DATE: 1998-10-19

PRIOR FILING DATE: 1995-06-07

NUMBER OF SEQ ID NOS: 16

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2

LENGTH: 139

TYPE: PRT

ORGANISM: Unknown

FEATURE:

OTHER INFORMATION: Description of Unknown Organism: EPIL - Early

OTHER INFORMATION: Placenta Insulin-Like Peptide

JS-09-599-564A-2

Alignment Scores:

Pred. No.: 2,296-83 Length: 139
Score: 733.00 Matches: 138
Percent Similarity: 99.28% Conservative: 0
Best Local Similarity: 99.28% Mismatches: 1
Query Match: 94.70% Indels: 0
JB: 4 Gaps: 0

JS-09-518-842-1 (1-420) x US-09-599-564A-2 (1-139)

Y 1 ATGGCCAGCTGCTCCGGTCTCTATCTGCGCAGCAATCTGGCTGCTGAGCAACTCCTT 60
b 1 MetAlaSerLeuPheArgSerTyrLeuProAlaIleTrpLeuLeuSerGlnLeuLeu 20
Y 61 AGAGAAAGCTAGCAGCAGAGCTGAGGGGATGTGTCCTCCGATTGGAAACACTTGTCTG 120
b 21 ArgGluSerLeuAlaAlaGluLeuArgGlyCysGlyProArgPheGlyLysHisLeuLeu 40
Y 121 TCATATTGCCCATCGCTGAGAGACATTCACCACCACCCAGGAGGCTGCTGCTGAA 180
b 41 SerTyrCysProMetProGluLysThrPheThrThrProGlyGlyTrpLeuLeuGlu 60
Y 181 TCTGACGCTCCCAAGAAATGTCGTAACCTCCACCAACAAAGATGGAAGCCTTAGGT 240
b 61 SerGlyArgProLysGluMetValSerThrSerLysAsnLysAspGlyGlnAlaLeuGly 80
Y 241 ACGACATCAGAAATTCATTCTTAATTGTCCAGAGCTGAAGAAACCACTGCTGAAGGG 300
b 81 ThrThrSerGluPheIleProAsnLeuSerProGluLeuLysLysProLeuSerGluGly 100
Y 301 CAGCCATCATTCAGAAATATATCTTCCCGCAAAAGAGAGAGTGGACGTCACAGATT 360
b 101 GlnProSerLeuLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 120
Y 361 GATCCATCTGTTGTCAAGTAATTTGTACCATGGAACCTTCAGTTAAATTTATGTACA 417
b 121 AspProPheCysCysGluValIleCysAspAspGlyThrSerValLysLeuCysThr 139

RESULT 4

S-09-174-465D-2

Sequence 2, Application US/09174465D

Patent No. 6180364

GENERAL INFORMATION:

APPLICANT: KOMAN, Ahment

APPLICANT: CHASSIN, Dorine

APPLICANT: BELLET, Dominique

TITLE OF INVENTION: NEW PROTEIN CALLED EPIL/PLACENTIN, PROCESS FOR THE

TITLE OF INVENTION: PREPARATION OF THIS PROTEIN AND PHARMACEUTICAL

TITLE OF INVENTION: COMPOSITION CONTAINING SUCH, DNA CODING FOR SAID

TITLE OF INVENTION: PROTEIN

FILE REFERENCE: 017753-103

CURRENT APPLICATION NUMBER: US/09/174,465D

CURRENT FILING DATE: 1998-10-19

PRIOR APPLICATION NUMBER: US 08/482,842

PRIOR FILING DATE: 1995-06-07

NUMBER OF SEQ ID NOS: 16

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2

LENGTH: 174

TYPE: PRT

ORGANISM: Unknown

FEATURE:

OTHER INFORMATION: Description of Unknown Organism: EPIL - Early

OTHER INFORMATION: Placenta Insulin-Like Peptide

US-09-174-465D-2

Alignment Scores:

Pred. No.: 2,566-83 Length: 174
Score: 733.00 Matches: 138
Percent Similarity: 99.28% Conservative: 0
Best Local Similarity: 99.28% Mismatches: 1
Query Match: 94.70% Indels: 0
DB: 3 Gaps: 0

US-09-518-842-1 (1-420) x US-09-174-465D-2 (1-174)

QY 1 ATGGCCAGCTGCTCCGGTCTCTATCTGCGCAGCAATCTGGCTGCTGAGCAACTCCTT 60
Db 36 MetAlaSerLeuPheArgSerTyrLeuProAlaIleTrpLeuLeuSerGlnLeuLeu 55
QY 61 AGAGAAAGCTAGCAGCAGAGCTGAGGGGATGTGTCCTCCGATTGGAAACACTTGTCTG 120
Db 56 ArgGluSerLeuAlaAlaGluLeuArgGlyCysGlyProArgPheGlyLysHisLeuLeu 75
QY 121 TCATATTGCCCATCGCTGAGAGACATTCACCACCACCCAGGAGGCTGCTGCTGAA 180
Db 76 SerTyrCysProMetProGluLysThrPheThrThrProGlyGlyTrpLeuLeuGlu 95
QY 181 TCTGACGCTCCCAAGAAATGTCGTAACCTCCACCAACAAAGATGGAAGCCTTAGGT 240
Db 96 SerGlyArgProLysGluMetValSerThrSerLysAsnLysAspGlyGlnAlaLeuGly 115
QY 241 ACGACATCAGAAATTCATTCTTAATTGTCCAGAGCTGAAGAAACCACTGCTGAAGGG 300
Db 116 ThrThrSerGluPheIleProAsnLeuSerProGluLeuLysLysProLeuSerGluGly 135
QY 301 CAGCCATCATTCAGAAATATATCTTCCCGCAAAAGAGAGAGTGGACGTCACAGATT 360
Db 136 GlnProSerLeuLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 155
QY 361 GATCCATCTGTTGTCAAGTAATTTGTACCATGGAACCTTCAGTTAAATTTATGTACA 417
Db 156 AspProPheCysCysGluValIleCysAspAspGlyThrSerValLysLeuCysThr 174

RESULT 5

US-08-991-890-5

Sequence 5, Application US/08991890

Patent No. 6114307

GENERAL INFORMATION:

APPLICANT: Jaepers, Stephen R.

APPLICANT: Sprugel, Katherine H.

APPLICANT: Ren, Hong Ping

APPLICANT: Humes, Jacqueline M.

APPLICANT: Hoffman, Ross C.

APPLICANT: Conklin, Darrell C.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR

TITLE OF INVENTION: STIMULATING PANCREATIC ISLET CELL REGENERATION

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSES: ZymoGenetics, Inc.

STREET: 1201 Eastlake Avenue East

CITY: Seattle

STATE: WA

COUNTRY: USA

ZIP: 98102
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/991,890
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/033,003
FILING DATE: December 16, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Sawislak, Deborah A
REGISTRATION NUMBER: 37,438
REFERENCE/DOCKET NUMBER: 96-41
TELEPHONE: 206-442-6672
TELEFAX: 206-442-6678
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 124 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-991-890-5

Alignment Scores:
Pred. No.: 3,23e-69 Length: 124
Score: 620.00 Matches: 114
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.13% Mismatches: 0
Query Match: 80.10% Indels: 0
DB: 3 Gaps: 0

US-09-518-842-1 (1-420) x US-08-991-890-5 (1-124)

Qy 73 GCACGAGCTGAGGGATGTCCTCCCGATTGGAAACACCTTGTCTCATATTGCCCC 132
Db 10 SerAlaGluLeuArgGlyCyG6GlyProArgPheGlyLysHisLeuLeuSerTyrCysPro 29
Qy 133 ATGCTGAGAGACATTCACACACACCCAGAGGGTGGCTGGATCTGGACGTCCC 192
Db 30 MetProGluLysThrPheThrThrThrProGlyGlyTyrLeuLeuGluSerGlyArgPro 49
Qy 193 AAGAATGGTGTCAACCTCCCAACAAAGATGGAACAGCCTTAGGTACGACATCGAA 252
Db 50 LysGluMetValSerThrSerAsnAsnLysAspGlyGlnAlaLeuGlyThrThrSerGlu 69
Qy 253 TTCATTCTTAATTTGTCCACGAGCTGAAGAAACCACTGTCTGAAGGCGAGCCATCTTG 312
Db 70 PheLeProAsnLeuSerProGluLeuLysLysProLeuSerGluGlyGlnProSerLeu 89
Qy 313 AAGAAATTAATCTTCCCGCAAAAGAGAGTGGACGTCAAGATTGTGATTCCTGT 372
Db 90 LysLysIleLeuSerArgLysLysArgSerGlyArgHisArgPheAspProPheCys 109
Qy 373 TGTGAAGTAATTTGTGACGATGGAACCTTCAGTTAAATATGTACA 417
Db 110 CysGluValIleCysAspAspGlyThrSerValLysLeuCysThr 124

RESULT 6

US-08-991-890-4
Sequence 4, Application US/08991890
Patent No. 6114307
GENERAL INFORMATION:
APPLICANT: Jaspers, Stephen R.
APPLICANT: Sprugel, Katherine H.
APPLICANT: Ren, Hong Ping
APPLICANT: Humes, Jacqueline M.
APPLICANT: Hoffman, Ross C.

APPLICANT: Conklin, Darrell C.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
TITLE OF INVENTION: STIMULATING PANCREATIC ISLET CELL REGENERATION
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: ZymoGenetics, Inc.
STREET: 1201 Eastlake Avenue East
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98102
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/991,890
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/033,003
FILING DATE: December 16, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Sawislak, Deborah A
REGISTRATION NUMBER: 37,438
REFERENCE/DOCKET NUMBER: 96-41
TELEPHONE: 206-442-6672
TELEFAX: 206-442-6678
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 159 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-991-890-4

Alignment Scores:
Pred. No.: 4,86e-69 Length: 159
Score: 619.00 Matches: 114
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 79.97% Indels: 0
DB: 3 Gaps: 0

US-09-518-842-1 (1-420) x US-08-991-890-4 (1-159)

Qy 76 GCACGCTGAGGGATGTCCTCCCGATTGGAAACACCTTGTCTCATATTGCCCATG 135
Db 46 AlaGluLeuArgGlyCysGlyProArgPheGlyLysHisLeuLeuSerTyrCysProMet 65
Qy 136 CTTGAGAGACATTCACACACCCAGAGGGTGGCTGCTGGAATCTCGAGTCCCAAA 195
Db 66 ProGluLysThrPheThrThrThrProGlyGlyTyrLeuLeuGluSerGlyArgProLys 85
Qy 196 GAAATGGTCAACCTCCCAACAAAGATGGAACCACTGTCTGAAGGCGAGCCATTCAG 255
Db 86 GluMetValSerThrSerAsnAsnLysAspGlyGlnAlaLeuGlyThrThrSerGluPhe 105
Qy 256 ATTCTTAATTTGTCCCGCAAAAGAGAGTGGACGTCAAGATTGTGATTCCTTGTGT 315
Db 106 IleProAsnLeuSerProGluLeuLysLysProLeuSerGluGlyGlnProSerLeuLys 125
Qy 316 AAAATAATTAATTTCCCGCAAAAGAGAGTGGACGTCAAGATTGTGATTCCTTGTGT 375
Db 126 LysIleLeuSerArgLysLysArgSerGlyArgHisArgPheAspProPheCys 145
Qy 376 GAAATAATTTGTGACGATGGAACCTTCAGTTAAATATGTACA 417

Db 146 GluValIleCysAspGlyThrSerValLysLeuCysThr 159

RESULT 7

US-09-174-465D-10
Sequence 10, Application US/09174465D
Patent No. 6180364

GENERAL INFORMATION:

APPLICANT: KOMAN, Ahment
APPLICANT: CHASSIN, Dorine
TITLE OF INVENTION: NEW PROTEIN CALLED EPIL/PLACENTIN, PROCESS FOR THE
TITLE OF INVENTION: PREPARATION OF THIS PROTEIN AND PHARMACEUTICAL
TITLE OF INVENTION: COMPOSITION CONTAINING SUCH, DNA CODING FOR SAID
TITLE OF INVENTION: PROTEIN

FILE REFERENCE: 017753-103

CURRENT APPLICATION NUMBER: US/09/174,465D

CURRENT FILING DATE: 1998-10-19

PRIOR APPLICATION NUMBER: US 08/482,842

PRIOR FILING DATE: 1995-06-07

NUMBER OF SEQ ID NOS: 16

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 10

LENGTH: 51

TYPE: PRT

ORGANISM: Unknown

FEATURE:

OTHER INFORMATION: Description of Unknown Organism: EPIL - Early

OTHER INFORMATION: Placenta Insulin-Like Peptide

US-09-174-465D-10

Alignment Scores:

Pred. No.:	1,27e-23	Length:	51
Score:	255.00	Matches:	51
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	32.95%	Indels:	0
DB:	3	Gaps:	0

US-09-518-842-1 (1-420) x US-09-174-465D-10 (1-51)

QY 175 CTGGAATCTGGACGTCCTCCAAAGAAATGGTGCAACTCCACACAAAGATGGACAGCC 234

Db 1 LeuGluSerGlyArgProLysGluMetValSerThrSerAsnAsnLysAspGlyGlnAla 20

QY 235 TTAGGTACGACATCAGATTCTTCTTAATTTGTCCACGAGCTGAAGAACCACTGTCT 294

Db 21 LeuGlyThrThrSerGluPheIleProAsnLeuSerProGluLeuLysLysProLeuSer 40

QY 295 GAAGGGCAGCCATCATTTGAAGAAAATAATCTT 327

Db 41 GluGlyGlnProSerLeuLysLysIleLeu 51

RESULT 8

US-09-599-564A-10

Sequence 10, Application US/09599564A

Patent No. 6362318

GENERAL INFORMATION:

APPLICANT: KOMAN, Ahment
APPLICANT: CHASSIN, Dorine
TITLE OF INVENTION: NEW PROTEIN CALLED EPIL/PLACENTIN, PROCESS FOR THE
TITLE OF INVENTION: PREPARATION OF THIS PROTEIN AND PHARMACEUTICAL
TITLE OF INVENTION: COMPOSITION CONTAINING SUCH, DNA CODING FOR SAID
TITLE OF INVENTION: PROTEIN

FILE REFERENCE: 017753-127

CURRENT APPLICATION NUMBER: US/09/599,564A

CURRENT FILING DATE: 2000-06-23

PRIOR APPLICATION NUMBER: 09/174,465

PRIOR FILING DATE: 1998-10-19

PRIOR APPLICATION NUMBER: US 08/482,842

PRIOR FILING DATE: 1995-06-07

NUMBER OF SEQ ID NOS: 16

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 10
LENGTH: 51
TYPE: PRT
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: EPIL - Early
OTHER INFORMATION: Placenta Insulin-Like Peptide
US-09-599-564A-10

Alignment Scores:

Pred. No.:	1,27e-23	Length:	51
Score:	255.00	Matches:	51
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	32.95%	Indels:	0
DB:	4	Gaps:	0

US-09-518-842-1 (1-420) x US-09-599-564A-10 (1-51)

QY 175 CTGGAATCTGGACGTCCTCCAAAGAAATGGTGCAACTCCACACAAAGATGGACAGCC 234

Db 1 LeuGluSerGlyArgProLysGluMetValSerThrSerAsnAsnLysAspGlyGlnAla 20

QY 235 TTAGGTACGACATCAGATTCTTCTTAATTTGTCCACGAGCTGAAGAACCACTGTCT 294

Db 21 LeuGlyThrThrSerGluPheIleProAsnLeuSerProGluLeuLysLysProLeuSer 40

QY 295 GAAGGGCAGCCATCATTTGAAGAAAATAATCTT 327

Db 41 GluGlyGlnProSerLeuLysLysIleLeu 51

RESULT 9

US-09-174-465D-8

Sequence 8, Application US/09174465D

Patent No. 6180364

GENERAL INFORMATION:

APPLICANT: KOMAN, Ahment
APPLICANT: CHASSIN, Dorine
APPLICANT: BELLET, Dominique
TITLE OF INVENTION: NEW PROTEIN CALLED EPIL/PLACENTIN, PROCESS FOR THE
TITLE OF INVENTION: PREPARATION OF THIS PROTEIN AND PHARMACEUTICAL
TITLE OF INVENTION: COMPOSITION CONTAINING SUCH, DNA CODING FOR SAID
TITLE OF INVENTION: PROTEIN

FILE REFERENCE: 017753-103

CURRENT APPLICATION NUMBER: US/09/174,465D

CURRENT FILING DATE: 1998-10-19

PRIOR APPLICATION NUMBER: US 08/482,842

PRIOR FILING DATE: 1995-06-07

NUMBER OF SEQ ID NOS: 16

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 8

LENGTH: 41

TYPE: PRT

ORGANISM: Unknown

FEATURE:

OTHER INFORMATION: Description of Unknown Organism: EPIL - Early

OTHER INFORMATION: Placenta Insulin-Like Peptide

US-09-174-465D-8

Alignment Scores:

Pred. No.:	2,77e-20	Length:	41
Score:	228.00	Matches:	41
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	29.46%	Indels:	0
DB:	3	Gaps:	0

US-09-518-842-1 (1-420) x US-09-174-465D-8 (1-41)

QY 52 CAACTCTTAGAAGAGCCCTAGCAGCAGCTGAGGGATGTGTCCTCCCGATTTCGAAA 111

Db 1 GlnLeuLeuArgGluSerLeuAlaGluLeuArgGlyCysGlyProArgPheGlyLys 20

Db 21 AspArgGlyThrSerValLeuCysThr 30
|||||

RESULT 13

US-09-599-564A-12

Sequence 12, Application US/09599564A

Patent No. 6362318

GENERAL INFORMATION:

APPLICANT: KOMAN, Ahment

APPLICANT: CHASSIN, Dorine

APPLICANT: BELLET, Dominique

TITLE OF INVENTION: NEW PROTEIN CALLED EPIL/PLACENTIN, PROCESS FOR THE

TITLE OF INVENTION: PREPARATION OF THIS PROTEIN AND PHARMACEUTICAL

TITLE OF INVENTION: COMPOSITION CONTAINING SUCH, DNA CODING FOR SAID

TITLE OF INVENTION: PROTEIN

FILE REFERENCE: 017753-127

CURRENT APPLICATION NUMBER: US/09/599,564A

CURRENT FILING DATE: 2000-06-23

PRIOR APPLICATION NUMBER: 09/174,465

PRIOR FILING DATE: 1998-10-19

PRIOR APPLICATION NUMBER: US 08/482,842

PRIOR FILING DATE: 1995-06-07

NUMBER OF SEQ ID NOS: 16

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 12

LENGTH: 30

TYPE: PRT

ORGANISM: Unknown

FEATURE:

OTHER INFORMATION: Description of Unknown Organism: EPIL - Early

OTHER INFORMATION: Placenta Insulin-Like Peptide

S-09-599-564A-12

Alignment Scores:

red. No.: 3,36e-13 Length: 30

core: 171.00 Matches: 30

percent Similarity: 100.00% Conservative: 0

est Local Similarity: 100.00% Mismatches: 0

uery Match: 22.09% Indels: 0

B: 4 Gaps: 0

S-09-518-842-1 (1-420) x US-09-599-564A-12 (1-30)

Y 328 TCCCGCAAAAGAGAGTGCACAGATTTCATTCCTGTTGAAGTAATTTGT 387

b 1 SerArgLysLysArgSerGlyArgHisArgPheAspPropheCysCysGluValIleCys 20

Y 398 GACGATGCAACTTCAGTTAAATTATGTACA 417

b 21 AspArgGlyThrSerValLeuCysThr 30

|||||

RESULT 14

S-09-201-227A-28

Sequence 28, Application US/09201227A

Patent No. 6468770

GENERAL INFORMATION:

APPLICANT: Keyes, Linda N.

APPLICANT: Doberstein, Stephen K.

APPLICANT: Buchman, Andrew R.

TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF D. MELANOGASTER INSULIN-LIKE

TITLE OF INVENTION: GENES AND USES THEREOF

FILE REFERENCE: 7326-066

CURRENT APPLICATION NUMBER: US/09/201,227A

CURRENT FILING DATE: 1998-11-30

NUMBER OF SEQ ID NOS: 45

SOFTWARE: PatentIn version 3.0

SEQ ID NO 28

LENGTH: 25

TYPE: PRT

ORGANISM: Homo sapiens

S-09-201-227A-28

Alignment Scores:

red. No.: 7,33e-07 Length: 185

Score: 123.50 Matches: 53

Percent Similarity: 35.35% Conservative: 17

Best Local Similarity: 26.77% Mismatches: 55

Query Match: 15.96% Indels: 73

DB: 3 Gaps: 7

Pred. No.: 3,15e-10 Length: 25

Score: 147.00 Matches: 25

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 18.99% Indels: 0

DB: 4 Gaps: 0

US-09-518-842-1 (1-420) x US-09-201-227A-28 (1-25)

QY 343 ACTGACGCTCAGATTTCATTCCTGTTGAAGTAATTTGTGACGATGGAACCTTCA 402

Db 1 SerGlyArgHisArgPheAspPropheCysCysGluValIleCysAspArgGlyThrSer 20

|||||

QY 403 GTTAAATTATGTACA 417

Db 21 ValLysLeuCysThr 25

|||||

RESULT 15

US-08-950-720A-12

; Sequence 12, Application US/08950720A

; Patent No. 6046028

; GENERAL INFORMATION:

; APPLICANT: Conklin, Darrell C.

; APPLICANT: Lofton-Day, Catherine E.

; APPLICANT: Lok, Si

; APPLICANT: Jaspers, Stephen R.

; TITLE OF INVENTION: INSULIN HOMOLOG

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: ZymoGenetics, Inc.

; STREET: 1201 Eastlake Avenue East

; CITY: Seattle

; STATE: WA

; COUNTRY: USA

; ZIP: 98102

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: Fast-Seq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/950,720A

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Sawislak, Deborah A.

; REGISTRATION NUMBER: 37,438

; REFERENCE/DOCKET NUMBER: 96-09

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 206-442-6672

; TELEFAX: 206-442-6678

; TELEX:

; INFORMATION FOR SEQ ID NO: 12:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 185 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: No. 6046028e

US-08-950-720A-12

Alignment Scores:

Pred. No.: 7,33e-07 Length: 185

Score: 123.50 Matches: 53

Percent Similarity: 35.35% Conservative: 17

Best Local Similarity: 26.77% Mismatches: 55

Query Match: 15.96% Indels: 73

DB: 3 Gaps: 7

US-09-518-842-1 (1-420) x US-08-950-720A-12 (1-185)


```

>Y 1 ATGGCCAGCTGTTCCGGTCTCTATCTGCACGAATCTGGTGTGCTGTGAGCAACTCCTT 60
>b ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
>Y 1 MetProArgLeuPhePhePheHisLeuLeuGlyValCysLeuLeuLeuAsnGlnPheSer 20
>b ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
>Y 61 -----ACAGAAAGCCTA 72
>b 21 ArgAlaValAlaAspSerTrpMetGluGluValIleLysLeuCysGlyArgGluLeuVal 40
>Y 73 GCACGAGAGCTGAGGGGATGTGGT---CCCGATTTCGAAACACACTTGTCTGATATTGC 129
>b ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
>Y 41 ArgAlaGlnIleAlaIleCysGlyMetSerThrTrpSerLysArgSerLeuSerGlnGlu 60
>b ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
>Y 130 CCCATGCTGTGAGAAACATTCCACCACCCAGGAGGGTGGTCTGGAATCTGGACGT 189
>b ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
>Y 61 AspAlaProGln-----ThrPro-----Arg 67
>b 190 CCC---AAGAAATGGTGTCAACCTCCAAACAACAAAGATGGACAGCCTTAGGTACGACA 246
>Y ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
>b 68 ProValAlaGluIleValProSerPheIleAsnLysAspThrGluThrIleAsnMetMet 87
>Y 247 TCAGAAATTCATTCTTAATTTGTCCACGAGCTGAAGAAACCACTGTCTGAAGGGCAGCCA 306
>b ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
>b 88 SerGluPheValAlaAsnLeuProGlnGluLeuLysLeuThrLeuSerGluMetGlnPro 107
>Y 307 TCATTG-----312
>b ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
>b 108 AlaLeuProGlnLeuGlnGlnHisValProValLeuLysAspSerSerLeuLeuPheGlu 127
>Y 313 -----AAGAAATTAATACTT-----327
>b ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
>b 128 GluPheLysLysLeuIleArgAsnArgGlnSerGluAlaAlaAspSerSerProSerGlu 147
>Y 328 -----TCCCGCAAAAGAGAGTGGACGTACAGATTT 360
>b ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
>b 148 LeuLysTyrLeuGlyLeuAspThrHisSerArgLysLysArgGlnLeuTyrSerAlaLeu 167
>Y 361 GATCCATTCTGTGTGAAGTAATTTGTGACGATGGAACCTTCAGTTAATTATGT 414
>b ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
>b 168 AlaAsnLysCysCysHisValGlyCysThrLysArgSerLeuAlaArgPheCys 185
```

search completed: October 9, 2003, 12:20:32
ob time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

M nucleic - protein search, using frame_plus_n2p model

on on: October 9, 2003, 12:09:01 ; Search time 16.5 Seconds
(without alignments)

2394.087 Million cell updates/sec

itle: US-09-518-842-1

Perfect score: 774

sequence:

1 ATGGCCAGCCTGTTCCGGT.....CACTAAATATGTCATAG 420

oring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 255726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

ommand line parameters:

MODEL=frame+ n2p.model -DEV=xlp
Q=/cgn2_1/USFO.spool/US09518842/runat_09102003_111035_13047/app query.fasta 1.583
DB=SwissProt_41 -OPMT=fastan -SUPFIX=rsp -MINMATCH=0.1 -DOOPCL=0 -DOOEXT=0
UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
DOCALIGN=200 -THR SCORES=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
USER=US09518842@cgn 1 1 30 @runat_09102003_111035_13047 -NCPU=6 -ICPU=3
NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPOP=0.5 -FGAPOP=6
FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

atabase : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

result	No.	Score	Query	Length	DB	ID	Description
1	739	95.5	139	1	INL4 HUMAN	Q14641	homo sapien
2	180	23.3	182	1	RELX_PIG	P01348	sus scrofa
3	151	19.5	185	1	RELI_MOUSE	P47932	mus musculus
4	147.5	19.1	182	1	RELX_HORSE	P22969	equus caball
5	123.5	16.0	185	1	REL2_HUMAN	P04090	homo sapien
6	122.5	15.8	185	1	RELX_MACMU	P19884	macaca mula
7	120	15.5	186	1	RELX_RAT	P01347	rattus norv
8	119.5	15.4	185	1	RELI_HUMAN	P04808	homo sapien
9	113.5	14.7	166	1	REL2_PANTR	P51455	pan troglod
10	111.5	14.4	178	1	RELH_RABIT	P51456	oryctolagus
11	108.5	14.0	166	1	RELH_PANTR	P51455	pan troglod
12	104.5	13.5	177	1	RELX_CANFA	Q64171	mesocricetu
13	95.5	12.3	177	1	RELX_MESAU	P24592	homo sapien
14	80	10.3	240	1	19P6_HUMAN	P24592	homo sapien
15	80	10.3	447	1	NUOP_BUCAP	Q8K9V3	buchnera ap
16	78.5	10.1	890	1	LNPI_HUMAN	Q14693	homo sapien
17	76.5	9.9	160	1	RELX_CAVPO	P51453	cavia porce
18	75	9.7	224	1	XP4_XENLA	Q00223	xenopus lae

19	75	9.7	836	1	DPOL_HPBOD	P03162	duck hepati
C 20	75	10.1	993	1	RPNI_YEAST	P38764	saccharomyc
C 21	75	10.1	1711	1	PTPO_RAT	Q64612	rattus norv
C 22	74.5	10.0	564	1	FEPI_SCHPO	Q10134	schizosacch
C 23	74.5	9.6	1531	1	PFPD_CHLTR	O84818	chlamydia t
C 24	74	9.9	857	1	AF56_CANAL	P78586	c argy, 6 pr
C 25	73.5	9.5	273	1	NADE_STAAM	P099x5	staphylococ
C 26	73.5	9.9	296	1	PERX_BRARA	P00434	brassica ra
C 27	73	9.4	1395	1	SP41_YEAST	P38904	saccharomyc
C 28	73	9.8	1744	1	CO4_HUMAN	P01028	homo sapien
C 29	72.5	9.4	292	1	YHFR_SALTI	Q824M8	salmonella
C 30	72.5	9.4	467	1	IL6A_PIG	O18796	sus scrofa
C 31	72	9.7	403	1	METM_CAEEL	O17680	caenorhabdi
C 32	72	9.3	752	1	YUJ3_SCHPO	Q9urt2	schizosacch
C 33	72	9.3	3358	1	PGCV_MOUSE	O62059	mus musculu
C 34	71.5	9.2	292	1	YFHR_SALTY	Q8zn39	salmonella
C 35	71.5	9.6	1021	1	PSKR_DAUCA	Q8lpB4	daucus caro
C 36	71	9.2	653	1	ACSA_SYNY3	Q55404	synecocyst
C 37	71	9.2	957	1	UVRA_BACSU	O34863	bacillus su
C 38	70.5	9.1	271	1	VG18_BPP22	P03687	bacterioph
C 39	70.5	9.5	1704	1	VILD_DICDI	Q8wg85	dictyosteli
C 40	70	9.0	283	1	PANC_SCHPO	Q09673	schizosacch
C 41	70	9.0	344	1	KLP3_MOUSE	Q60980	mus musculu
C 42	70	9.4	388	1	METR_ACACA	Q95032	acanthamoeb
C 43	70	9.4	486	1	CDSN_HUMAN	Q15517	homo sapien
C 44	70	9.4	670	1	YFGG_SCHPO	O13854	schizosacch
C 45	70	9.4	1356	1	ROM2_YEAST	P51862	saccharomyc

ALIGNMENTS

RESULT 1	INL4_HUMAN	STANDARD;	PRT; 139 AA.
ID	INL4_HUMAN	STANDARD;	PRT; 139 AA.
AC	Q14641;		
DT	01-NOV-1997 (Rel. 35, Created)		
DT	01-NOV-1997 (Rel. 35, Last sequence update)		
DT	15-SEP-2003 (Rel. 42, Last annotation update)		
DE	Early placenta insulin-like peptide precursor (EPIL) (Placental)		
DE	(Insulin-like peptide 4).		
GN	INSL4.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Placenta;		
RX	MEDLINE=96115599; PubMed=8666396;		
RA	Chassin D., Laurent A., Janneau J.-L., Berger R., Bellet D.;		
RT	"Cloning of a new member of the insulin gene superfamily (INSL4)		
RT	expressed in human placenta.";		
RL	Genomics 29:465-470(1995).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Placenta;		
RX	MEDLINE=22388257; PubMed=12477932;		
RA	Strausberg R.L., Peingold E.A., Grouse L.H., Derge J.G.,		
RA	Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,		
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,		
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,		
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,		
RA	Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,		
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,		
RA	Raha S., Loquellano N.A., McKernan K.J., Abramson R.D., Mullaly S.J.,		
RA	Bosak S.A., McQuinn P.J., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,		
RA	Villalón D.K., Murzy D.M., Sodergren E.J., Lu X., Gibbs R.A.,		
RA	Fahey J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,		
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,		
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,		
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,		
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,		

SEQUENCE OF 25-51.
MEDLINE=715271; PubMed=851452;
Schwabe C., McDonald J.K., Steinetz B.G.;
"Primary structure of the B-chain of porcine relaxin."
Biochem. Biophys. Res. Commun. 75:503-510(1977).
[5]
SEQUENCE OF 25.


```
T DISULFID 35 169 INTERCHAIN (BY SIMILARITY).
T DISULFID 47 182 INTERCHAIN (BY SIMILARITY).
T DISULFID 168 173 BY SIMILARITY.
T CONFLICT 66 66 A -> V (IN REF. 2).
T CONFLICT 133 133 L -> Q (IN REF. 2).
Q SEQUENCE 182 AA; 20721 MW; 55C9414303A838B8 CRC64;

Ligament Scores:
red. No.: 1.09e-07 Length: 182
core: 147.50 Matches: 43
Percent Similarity: 52.59% Conservative: 18
Best Local Similarity: 37.07% Mismatches: 44
Query Match: 19.06% Indels: 11
JB: 1 Gaps: 4

US-09-518-842-1 (1-420) x RELX_HORSE (1-182)
Y 1 ATGCCAGCGCTGTCGGCTCTATCTCCAGCAATCTGGCTGCTGAGCCAACTCCTT 60
b 1 MetArgLeuPheLeuSerHisValLeuGlyAlaTrpLeuLeuLeuSerGlnLeuPro 20
Y 61 AGACAAAGCCCTAGCAGCAGAG-----CTGAGGGGATGCTGCTCCCGATTGGA 108
b 21 ArgGluLeuSerGlyGlnLysProAspValIleLysAlaCysGlyArgGluLeuAla 40
Y 109 AAACACTTGCTCATATTGCCCATGCT-----GAGAAGACATTC-----ACCACC 156
b 41 ArgLeuArgIleGluIleCysGlySerLeuSerTrpLysLysThrValLeuArgLeuGlu 60
Y 157 ACCCAGAGAGGTGCTGCTGGAATCTGGAGTCCCAAGAAATGTTGCTCAACTCAAC 216
b 61 GluProGly-----LeuGluAlaGlyGlnProValGluIleValSerSerIle 77
Y 217 AACAAAGATGACAGCGCTAGGTACCATCATTCATTCCTCAATTGTCACCCAGAG 276
b 78 SerLysAspAlaGluAlaLeuAsnThrLysLeuGlyLeuAsnSerAsnLeuProLysGlu 97
Y 277 CTGAAGAAACCACTGCTCTGAGGCGCAGCCATCATGTAAGAAATAATA 324
b 98 GlnLysAlaThrLeuSerGluArgGlnProSerTrpArgGluLeuLeu 113

RESULT 5
EL2 HUMAN STANDARD; Q93UJ2; PRT; 185 AA.
D REL2 HUMAN STANDARD; Q93UJ2; PRT; 185 AA.
C P04090; Q99936; Q93UJ2;
T 01-NOV-1986 (Rel. 03, Created)
T 01-NOV-1986 (Rel. 03, Last sequence update)
T 15-SEP-2003 (Rel. 42, Last annotation update)
E Prorelaxin H2 precursor.
N Homo sapiens (Human).
S Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
C Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
X NCBI_TaxID=9606;
P (1)
N SEQUENCE FROM N.A. (ISOFORM 1), AND TISSUE SPECIFICITY.
X MEDLINE=85051296; PubMed=6548702;
X Hudson P., John M., Crawford R., Haralambidis J., Scanlon D.,
A Gorman J., Tregear G., Shine J., Niall H.;
T "Relaxin gene expression in human ovaries and the predicted structure
of a human preprorelaxin by analysis of cDNA clones."
L EMBO J. 3:2333-2339(1984).
P (2)
N SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
X Sehra H.;
L Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
P (3)
N PARTIAL SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE
SPECIFICITY.
P TISSUE=Prostate;
X MEDLINE=96328899; PubMed=8735594;
X Gunnersen J.M., Fu P., Roche P.J., Tregear G.W.;
T "Expression of human relaxin genes: Characterization of a novel
```

```
RT alternatively-spliced human relaxin mRNA species.";
RL Mol. Cell. Endocrinol. 118:85-94(1996).
RN (4).
RP SEQUENCE OF 1-15 FROM N.A.
RA Garibay-Tupas J.;
RT "Characterization of the human relaxins H1 and H2 5'-flanking
regions.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN (5).
RP PARTIAL SEQUENCE OF 25-51.
RC TISSUE=Semen;
RX MEDLINE=92241162; PubMed=1572287;
RA Winlow J.W., Shih A., Bourrell J.H., Weiss G., Reed B., Stultes J.T.,
RA Goldsmith L.T.;
RT "Human seminal relaxin is a product of the same gene as human luteal
relaxin.";
RL Endocrinology 130:2660-2668(1992).
RN (6).
RP SEQUENCE OF 25-53 AND 162-185.
RX MEDLINE=91167739; PubMed=2076464;
RA Winlow J.W., Griffin P.R., Rinderknecht E., Vandlen R.L.;
RT "Structural characterization by mass spectrometry of native and
recombinant human relaxin.";
RL Biomed. Environ. Mass Spectrom. 19:655-664(1990).
RN (7).
RP SYNTHESIS.
RX MEDLINE=91250367; PubMed=2040595;
RA Buellesbach E.E., Schwabe C.;
RT "Total synthesis of human relaxin and human relaxin derivatives by
solid-phase peptide synthesis and site-directed chain combination.";
RL J. Biol. Chem. 266:10754-10761(1991).
RN (8).
RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).
RX MEDLINE=92015205; PubMed=1656049;
RA Eigenbrodt C., Randal M., Quan C., Burnier J., O'Connell L.,
RA Rinderknecht E., Kossiakoff A.A.;
RT "X-ray structure of human relaxin at 1.5 A. Comparison to insulin and
implications for receptor binding determinants.";
RL J. Mol. Biol. 221:15-21(1991).
CC -1- FUNCTION: RELAXIN IS AN OVARIAN HORMONE THAT ACTS WITH ESTROGEN TO
PRODUCE DILATATION OF THE BIRTH CANAL IN MANY MAMMALS. MAY BE
INVOLVED IN REMODELING OF CONNECTIVE TISSUES DURING PREGNANCY,
PROMOTING GROWTH OF PUBIC LIGAMENTS AND RIPENING OF THE CERVIX.
-1- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
DISULFIDE BONDS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=2;
Name=1;
Name=2;
IsoID=P04090-1; Sequence=Displayed;
Name=2;
IsoID=P04090-2; Sequence=VSP 002711, VSP 002712;
-1- TISSUE SPECIFICITY: Expressed in the ovary during pregnancy. Also
expressed in placenta, decidua and prostate.
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL: X00948; CAA25460.1; -.
DR EMBL: AL135786; CAC04177.1; -.
DR EMBL: AL135786; CAC04176.1; -.
DR EMBL: S83200; AAD14429.1; -.
DR EMBL: AF104935; AAD21961.1; -.
DR EMBL: AF1315; CAA01324.1; -.
DR EMBL: A06925; CAA00602.1; -.
DR PIR: A05092; A60982.
DR PDB: 6RLX; 31-OCT-93.
```

```

DR MIM; HGNC:10027; RIN2.
DR GO; GO:0007565; P:pregnancy; TAS.
DR InterPro; IPR004825; Ins/IGF/relax.
DR Pfam; PF00049; Insulin; 1.
DR SMART; SM00078; IIGF; 1.
DR PROSITE; PS00262; INSULIN; 1.
KW Insulin family; Hormone; Multigene family; Signal; 3D-structure;
KW Alternative splicing; Pyrrolidone carboxylic acid.
DR SIGNAL 1 24
FT CHAIN 25 53
FT PROPEP 56 157
FT CHAIN 162 185
FT MOD_RES 162 162
FT DISULFID 35 172
FT DISULFID 47 185
FT DISULFID 171 176
FT VARSPLIC 71 117
FT
FT CHAIN 25 53
FT PROPEP 56 157
FT CHAIN 162 185
FT MOD_RES 162 162
FT DISULFID 35 172
FT DISULFID 47 185
FT DISULFID 171 176
FT VARSPLIC 71 117
FT
FT VARSPLIC 118 185
FT
FT HELIX 27 29
FT STRAND 31 32
FT HELIX 36 49
FT HELIX 164 173
FT TURN 174 174
FT STRAND 176 177
FT HELIX 178 182
FT TURN 183 185
3Q *SEQUENCE 185 AA; 21042 MW; AC73DBDE2090091B CRC64;

Alignment Scores:
Pred. No.: 3,94E-05 Length: 185
Score: 123.50 Matches: 53
Percent Similarity: 35.34% Conservative: 17
Best Local Similarity: 26.77% Mismatches: 55
Query Match: 15.96% Indels: 73
DB: 1 Gaps: 7

US-09-518-842-1 (1-420) x REL2_HUMAN (1-185)
Y 1 ATGGCCAGCGCTCTCGGTCCTATCTGCCAGCAATCTGCTGCTGAGCAACTCCT 60
Y 1 MetProArgLeuPhePhePheHisLeuLeuGlyValCysLeuLeuLeuAsnGlnPheSer 20
Y 61 -----AGAGAAAGCCTA 72
Y 21 ArgAlaValAlaAspSerTrpMetGluGluValIleLysLeuCysGlyArgGluLeuVal 40
Y 73 GCAGCAGACGCTGAGGCGATGCTGCT--CCCGATTGGAAACACTCTGCTCATATTGC 129
Y 41 ArgAlaGlnIleAlaIleCysGlyMetSerThrTrpSerLysArgSerLeuSerGlnGlu 60
Y 130 CCATGCTGAGAGACATTCACCAACCCAGGAGGTGCTGCTGGAATCTGGACGT 189
Y 61 AspAlaProGln-----ThrPro-----Arg 67
Y 190 CCC--AAAGAAATGGTGTCACTCCACACACAAAGATGGACAGCTTAGGTACGACA 246
Y 68 ProValAlaGluIleValProSerPheIleAsnLysAspThrGluThrIleAsnMetMet 87
Y 247 TCAGAAATTCCTCTAATTGTCACACAGAGCTGGAAGAACCACTCTCTGAAGGGCAGCCA 306
Y 88 SerGluPheValAlaAsnLeuProGlnGluLeuLysLeuThrLeuSerGluMetGlnPro 107
Y 307 TCATTG-----
Y 108 AlaLeuProGlnLeuGlnGlnHisValProValLeuLysAspSerSerLeuLeuPheGlu 127
Y 313 -----AAGAAATATACCTT-----
Y 312 -----AGAGAAAGCCTA 72
Y 21 ArgAlaValAlaAlaLysTrpMetAspValIleLysAlaCysGlyArgGluLeuVal 40

```

```

Db 128 GluPheLysLysLeuIleArgAsnArgGlnSerGluAlaAlaAspSerSerProSerGlu 147
QY 328 -----TCCGCAAAAGAGAGAGTGGACGTGCACAGATTT 360
Db 148 LeuLysTyLeuGlyLeuAspThrHisSerArgLysArgGlnLeuTySerAlaLeu 167
QY 361 GATCAATCTCTGTGTGAAGTAATTGTGACGATGAACCTCAGTTAAATATGT 414
Db 168 AlaAsnLysCysHisValGlyCysThrLysArgSerLeuAlaArgPheCys 185

RESULT 6
RELX_MACMU
ID RELX_MACMU STANDARD; PRT; 185 AA.
AC P19884;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Prorelaxin precursor.
GN RIN.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90073957; PubMed=2590381;
RA Crawford R.J., Hammond V.E., Roche P.J., Johnston P.D., Tregear G.W.;
RT "Structure of rhesus monkey relaxin predicted by analysis of the
RT single-copy rhesus monkey relaxin gene.";
RL J. Mol. Endocrinol. 3:169-174(1989).
CC -!- FUNCTION: RELAXIN IS AN OVARIAN HORMONE THAT ACTS WITH ESTROGEN TO
CC PRODUCE DILATATION OF THE BIRTH CANAL IN MANY MAMMALS. MAY BE
CC INVOLVED IN REMODELING OF CONNECTIVE TISSUES DURING PREGNANCY.
CC PROMOTING GROWTH OF PUBIC LIGAMENTS AND RIPENING OF THE CERVIX.
CC -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR PIR; A34936; A34936.
DR HSP; P04090; 6RLX.
DR InterPro; IPR004825; Ins/IGF/relax.
DR Pfam; PF00049; Insulin; 1.
DR SMART; SM00078; IIGF; 1.
DR PROSITE; PS00262; INSULIN; 1.
KW Insulin family; Hormone; Signal.
FT SIGNAL 1 22
FT CHAIN 23 53
FT PROPEP 56 157
FT CHAIN 161 185
FT DISULFID 35 172
FT DISULFID 47 185
FT DISULFID 171 176
FT DISULFID 171 176
SQ SEQUENCE 185 AA; 20895 MW; 7E3C5D21B57E185C CRC64;

Alignment Scores:
Pred. No.: 5.03E-05 Length: 185
Score: 122.50 Matches: 51
Percent Similarity: 34.34% Conservative: 17
Best Local Similarity: 25.76% Mismatches: 57
Query Match: 15.83% Indels: 73
DB: 1 Gaps: 6

US-09-518-842-1 (1-420) x RELX_MACMU (1-185)
QY 1 ATGGCCAGCGCTCTCGGTCCTATCTGCCAGCAATCTGCTGCTGAGCAACTCCT 60
Db 1 MetProArgLeuPhePheHisLeuLeuGlyValCysLeuLeuLeuAsnGlnPheSer 20
QY 61 -----AGAGAAAGCCTA 72
Db 21 ArgAlaValAlaAlaLysTrpMetAspValIleLysAlaCysGlyArgGluLeuVal 40

```

```

73 GCAGCAGAGCTGGAGGATGTGGT---CCCGATTTGGAACACACTTGTCTCATATTGC 129
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
41 ArgAlaGlnIleAlaIleCysGlySerThrLeuGlyLysArgSerLeuAsnGlnGlu 60
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
130 CCATGCTGCTGAGAGACATTCACACCCAGGAGGGTGGTGTGGAATCTGGACGT 189
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 AspAlaProLeuLysPro-----Arg 67
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
190 CCC---AAGAAATGCTCACTCCCAACAAAGATGGACAGCCTTAGGTACGACA 246
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
68 ProAlaAlaGluIleValProSerLeuIleAsnGlnAspThrGluThrIleAsnMetMet 87
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
247 TCAGAAATTCATTCCTAATTTGTCCACAGAGCTGAAGAAACCACTGTCTGAAGGCGACCA 306
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
88 SerGluPheValAlaAsnLeuProGlnGluLeuLysLeuThrLeuSerGluArgGlnPro 107
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
307 TCATTG-----312
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
108 AlaLeuSerGluLeuGlnGlnHisValProValLeuLysAspSerAsnLeuSerPheGlu 127
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
313 -----AAGAAATATAATACIT-----327
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
128 GluPheLysLysIleIleArgLysArgGlnSerGluAlaThrAspSerProSerGlu 147
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
328 -----TCCGCAAAAGAGAAAGTGGACGTACAGATT 360
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
148 LeuArgSerLeuGlyLeuAspThrHisSerArgLysArgGlnLeuTyMetThrLeu 167
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
361 GATCATTCTGTTGTAAGTAATTTTGAGATGGAACTTCAGTTAAATATGT 414
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
168 SerAsnLysCysHisIleGlyCysThrLysLysSerLeuAlaLysPheCys 185
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

RESULT 7

```

ELX-RAT          STANDARD;          PRT;          186 AA.
D RELX-RAT
C P01347;
T 21-JUL-1986 (Rel. 01, Created)
T 21-JUL-1986 (Rel. 01, Last sequence update)
T -28-FEB-2003 (Rel. 41, Last annotation update)
E Prorelaxin precursor.
N RLN
S Rattus norvegicus (Rat).
C Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
C Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
X NCBI_TaxID=10116;
N [1]
S SEQUENCE FROM N.A.
MEDLINE=81197624; PubMed=7231533;
X A Hudson P., Haley J., Cronk M., Shine J., Niall H.;
T "Molecular cloning and characterization of cDNA sequences coding for
T rat relaxin.";
L Nature 291:127-131(1981).
N [2]
S SEQUENCE OF 23-57 AND 163-188.
MEDLINE=8109283; PubMed=7004862;
X A John M.J., Borjesson B.W., Walsh J.R., Niall H.D.;
T "Limited sequence homology between porcine and rat relaxins:
T implications for physiological studies.";
L Endocrinology 108:726-729(1981).
C -!- FUNCTION: RELAXIN IS AN OVARIAN HORMONE THAT ACTS WITH ESTROGEN
C TO PRODUCE DILATATION OF THE BIRTH CANAL IN MANY MAMMALS.
C -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
C DISULFIDE BONDS.
C -!- SUBCELLULAR LOCATION: Secreted.
C -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
C
C This SWISS-PROT entry is copyright. It is produced through a collaboration
C between the Swiss Institute of Bioinformatics and the EMBL outstation -
C the European Bioinformatics Institute. There are no restrictions on its
C use by non-profit institutions as long as its content is in no way
C modified and this statement is not removed. Usage by and for commercial
C entities requires a license agreement (See http://www.isb-sib.ch/announce/
C or send an email to license@isb-sib.ch).

```

```

CC EMBL: J00780; AAA42029.1; -
DR EMBL: V01264; CAA24578.1; -
DR FIR: A01614; RXRT.
DR HSSP: P01348; 1RLX.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF00049; Insulin; 1.
DR SMART: SM00078; ILGF; 1.
DR PROSITE: PS00262; INSULIN; 1.
KW Insulin family; Hormone; Signal; Pyrrolidone carboxylic acid.
FT SIGNAL 1 22
FT CHAIN 23 57 RELAXIN B CHAIN.
FT PROPEP 58 158 CONNECTING PEPTIDE.
FT CHAIN 163 186 RELAXIN A CHAIN.
FT MOD_RES 163 163 PYRROLIDONE CARBOXYLIC ACID.
FT DISULFID 36 173 INTERCHAIN (BY SIMILARITY).
FT DISULFID 48 186 INTERCHAIN (BY SIMILARITY).
FT DISULFID 172 177 BY SIMILARITY.
SQ SEQUENCE 186 AA; 20489 MW; 08BAEC79BCF0E80F CRC64;

Alignment Scores:
Pred. No.: 9.3e-05 Length: 186
Score: 120.00 Matches: 46
Percent Similarity: 33.16% Conservative: 18
Best Local Similarity: 23.83% Mismatches: 67
Query Match: 15.50% Indels: 62
DB: 1 Gaps: 5

US-09-518-842-1 (1-420) x RELX_RAT (1-186)
QY 1 ATGGCCAGCTGTTCCGGTCTCTATCTGCCAGCAATCTGGCTGCTGAGCAACTCCTT 60
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MetSerSerArgLeuLeuLeuGlnLeuLeuGlyPheTrpLeuPheLeuSerGlnProCys 20
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 AGAGAAACCTTAGCAGCAGAG-----CTGAGGGGATGTGTCCTCCCGATT 105
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 21 ArgAlaArgValSerGluGluTrpMetAspGlnValIleGlnValCysGlyArgGlyTyr 40
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 106 GCAAAACACTTGTCTCATATTGCCCCATGCTCAGAGACATTCACCCACCCACGGA 165
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 41 AlaArgAlaTrpIleGluValCys-----GlyAlaSerValGly 53
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 166 GGGTGGCTGCTGGAATCTGGACGTCCC-----AAGAAATG 201
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 54 ArgLeuAlaLeuSerGlnGluProAlaProLeuAlaArgGlnAlaThrAlaGluVal 73
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 202 GTGTCAACCTCCACACAAAGATGGACACCTTAGTCAGCATCAGAAATTCATTCCT 261
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 74 ValProSerPheIleAsnLysAspAlaGluProPheAspMetThrLeuLysCysLeuPro 93
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 262 AATTTGTCCACAGAGCTGAAGAAACCACTGTCTGAAGGGCAG-----303
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 94 AsnLeuSerGluGluArgLysAlaLeuSerGluGlyArgAlaProPheProGluLeu 113
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 303 -----303
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 114 GlnGlnHisAlaProAlaLeuSerAspSerValValSerLeuGluGlyPheLysLysThr 133
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 304 -----CCATCATTTGAAGAAATA---321
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 134 PheHisAsnGlnLeuGlyGluAlaGluAspGlyGlyProProGluLeuLysTyrLeuGly 153
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 322 -----ATACTTTCCCGCAAAAGAGAGTGCAGCTCACAGATTGATTCATTCGTGT 375
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 154 SerAspAlaGlnSerArgLysLysArgGlnSerGlyAlaLeuLeuSerGluGlnCysCys 173
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 376 GAAGTAATTTGTGAGAGTGAACCTTCAGTTAAATATGT 414
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 174 HisIleGlyCysThrArgArgSerIleAlaLysLeuCys 186
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

RESULT 8

```

RELX_HUMAN
ID RELX_HUMAN STANDARD; PRT; 185 AA.

```


P04808; Q99936; Q9UQ11;
 13-AUG-1987 (Rel. 05, Created)
 13-AUG-1987 (Rel. 05, Last sequence update)
 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Prelaxin H1 precursor.
 RLNI.
 CC Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 CC NCBI_TaxID=9606;
 CC [1]
 CC SEQUENCE FROM N.A. (ISOFORM 1).
 CC MEDLINE=85051298; PubMed=6548702;
 CC Hudson P., John M., Crawford R., Haralambidis J., Scanlon D.,
 CC Gorman J., Tregear G., Shine J., Niall H.;
 CC "Relaxin gene expression in human ovaries and the predicted structure
 CC of a human preprorelaxin by analysis of cDNA clones";
 CC EMBO J. 3:2333-2339(1984).
 CC [2]
 CC SEQUENCE FROM N.A. (ISOFORM 1).
 CC MEDLINE=83141755; PubMed=6298628;
 CC Hudson P., Haley J., John M., Cronk M., Crawford R., Haralambidis J.,
 CC Tregear G., Shine J., Niall H.;
 CC "Structure of a genomic clone encoding biologically active human
 CC relaxin";
 CC Nature 301:628-631(1983).
 CC [3]
 CC SEQUENCE FROM N.A. (ISOFORM 1).
 CC Submitted (AUG-2000) to the EMBL/GenBank/DBSJ databases.
 CC [4]
 CC SEQUENCE FROM N.A. (ISOFORM 1).
 CC TISSUE=Prostate;
 CC MEDLINE=22388257; PubMed=12477932;
 CC Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 CC Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 CC Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 CC Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 CC Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 CC Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 CC Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Frange C.,
 CC Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 CC Bosak S.A., McEwan P.C., McKernan K.J., Malek J.A., Gunaratne P.H.,
 CC Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 CC Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
 CC Fahy J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
 CC Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 CC Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 CC Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 CC Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalios D.E.,
 CC Schnerch A., Schein J.E., Jones S.J.M., Marita M.A.;
 CC "Generation and initial analysis of more than 15,000 full-length
 CC human and mouse cDNA sequences";
 CC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC [5]
 CC PARTIAL SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE
 CC SPECIFICITY.
 CC TISSUE=Prostate;
 CC MEDLINE=96328899; PubMed=8735594;
 CC Gunnarsen J.M., Fu P., Roche P.J., Tregear G.W.;
 CC "Expression of human relaxin genes: characterization of a novel
 CC alternatively-spliced human relaxin mRNA species";
 CC Mol. Cell. Endocrinol. 118:85-94(1996).
 CC [6]
 CC SEQUENCE OF 1-11 FROM N.A.
 CC Garibay-Tupas J.;
 CC "Characterization of the human H1 relaxin 5'-flanking region.";
 CC Submitted (NOV-1998) to the EMBL/GenBank/DBSJ databases.
 CC -1- FUNCTION: RELAXIN IS AN OVARIAN HORMONE THAT ACTS WITH ESTROGEN TO
 CC PRODUCE DILATATION OF THE BIRTH CANAL IN MANY MAMMALS. MAY BE
 CC INVOLVED IN REMODELING OF CONNECTIVE TISSUES DURING PREGNANCY,
 CC PROMOTING GROWTH OF PUBIC LIGAMENTS AND RIPENING OF THE CERVIX.
 CC -1- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO

CC DISULFIDE BONDS.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=P04808-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=P04808-2; Sequence=VSP_002709, VSP_002710;
 CC -1- TISSUE SPECIFICITY: Prostate. Not expressed in placenta, decidua
 CC or ovary.
 CC -1- MISCELLANEOUS: H1 RELAXIN MAY BE A PSEUDOCENE.
 CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@sib-sib.ch).
 CC EMBL; X00949; CAA25461.1; -;
 CC EMBL; X00578; CAA23839.1; -;
 CC EMBL; V00577; CAA23838.1; -;
 CC EMBL; AF104934; RAD21967.1; -;
 CC EMBL; A06926; CAA00603.1; -;
 CC EMBL; A06846; CAA00599.1; -;
 CC EMBL; A07364; CAA00658.1; -;
 CC EMBL; A17329; CAA01325.1; -;
 CC EMBL; AL135786; CAC04179.1; -;
 CC EMBL; BC005956; AAH05936.1; -;
 CC EMBL; S83200; AAD14429.1; -;
 CC FIR; B05092; A44559.
 CC HSSP; P04090; 6RLX.
 CC Genew; HGNC:10026; RLNI.
 CC MIM; 179730; -;
 CC GO; GO:0005180; F:peptide hormone; TAS.
 CC InterPro: IPR004825; Ins/IGF/relax.
 CC Pfam; PF00049; Insulin; 1.
 CC SMART; SM00078; IGF; 1.
 CC PROSITE; PS00262; INSULIN; 1.
 CC Insulin family; Hormone; Multigene family; Signal;
 CC Alternative splicing; Polymorphism.
 CC SIGNAL 1 22 PROBABLE.
 CC CHAIN 23 53 RELAXIN B CHAIN (PROBABLE).
 CC PROPEP 56 158 CONNECTING PEPTIDE (PROBABLE).
 CC CHAIN 163 185 RELAXIN A CHAIN (PROBABLE).
 CC DISULFID 35 172 INTERCHAIN (BY SIMILARITY).
 CC DISULFID 47 185 INTERCHAIN (BY SIMILARITY).
 CC DISULFID 171 176 BY SIMILARITY.
 CC VARSPLIC 71 117 EIVPSFINKDTETIIIMLEFIANLPPELKAALSERQPSLPE
 CC LQQYVP -> GDFIOTVSLGISPDGKALRTGSCFTREFLG
 CC ALSKLYHPSSTKIOKL (in isoform 2).
 CC /FTID=VSPF_002709.
 CC Missing (in isoform 2).
 CC /FTID=VSP_002710.
 CC K -> M (in dbSNP:618066).
 CC /FTID=VAR_011962.
 CC SQ SEQUENCE 185 AA; 21145 MW; B318628ABFEC7142 CRC64;
 Alignment Scores:
 Pred. No.: 0.000105 Length: 185
 Score: 119.50 Matches: 52
 Percent Similarity: 37.11% Conservative: 20
 Best Local Similarity: 26.80% Mismatches: 57
 Query Match: 15.44% Indels: 65
 DB: 1 Gaps: 7
 US-09-518-842-1 (1-420) x RELI_HUMAN (1-185)
 QY 1 ATGCGACGCTGTTCGGGCTCTATCTGCAGCAATCTGGCTGCTGCTGAGCAACTCCTT 60
 Db 1 MetProArgLeuPheLeuPheHisLeuGluPheCysLeuLeuLeuGlnPheSer 20


```

313 -----AAGAAATAATACTT----- 327
105 LeuLeuPheGluGluPheLysLysLeuIleArgAsnArgInSerGluAlaAlaAaspSer 124
328 -----TCCCGCAAAAGAGAGTGA 348
125 SerProSerGluLeuLysTyrLeuGlyLeuAaspThrHisSerArgLysLysArgGlnLeu 144
349 COTCAGATTGATCCATCTGTTGTGAAGTAATTTGTGACGATGGAACCTTCAGTTAA 408
145 TyrSerAlaLeuAlaAsnLysCysShisValGlyCysThrLysArgSerLeuAlaArg 164
409 TTAATCT 414
165 PheCys 166

RESULT 10
RELH_RABBIT
ID RELH_RABBIT STANDARD; PRT; 178 AA.
AC P51456;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Relaxin-like protein SQ10 precursor.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RT "TISSUE=Tracheobronchial epithelium;
RX MEDLINE=93002619; PubMed=1339318;
RA Jettten A.M., Bernacki S.H., Floyd E.E., Saunders N.A., Pieniazek J.,
RA Lotan R.;
RA "Expression of a preprorelaxin-like gene during squamous
RT differentiation of rabbit tracheobronchial epithelial cells and its
RT suppression by retinoic acid."
RL Cell Growth Differ. 3:549-556(1992).
RC
CX -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CX DISULFIDE BONDS (BY SIMILARITY).
CX -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CX -!- INDUCTION: DURING SQUAMOUS CELL DIFFERENTIATION. REPRESSED BY
CX RETINOIC ACID.
CX -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CX
CX This SWISS-PROT entry is copyright. It is produced through a collaboration
CX between the Swiss Institute of Bioinformatics and the EMBL outstation -
CX the European Bioinformatics Institute. There are no restrictions on its
CX use by non-profit institutions as long as its content is in no way
CX modified and this statement is not removed. Usage by and for commercial
CX entities requires a license agreement (See http://www.ebi.ac.uk/ebis-sib.ch/announcement/
CX or send an email to license@sib-sib.ch).
CX
CX -----
CX EMBL; 545940; AAB23648.1; -.
CX PIR; A49014; A49014.
CX InterPro; IPR004825; Ins/IGF/relax.
CX Pfam; PF00049; Insulin; 1.
CX SMART; SM00078; IIGF; 1.
CX PROSITE; PS00262; INSULIN; 1.
CX Insulin family; Hormone; Signal.
CX SIGNAL 1 20
CX CHAIN 21 52
CX POTENTIAL.
CX RELAXIN-LIKE PROTEIN SQ10 B CHAIN
CX (POTENTIAL).
CX CONNECTING PEPTIDE (POTENTIAL).
CX RELAXIN-LIKE PROTEIN SQ10 A CHAIN
CX (POTENTIAL).
CX INTERCHAIN (BY SIMILARITY).
CX INTERCHAIN (BY SIMILARITY).
CX DISULFID 34 165
CX DISULFID 46 178
CX DISULFID 164 169
CX BY SIMILARITY.
CX SEQUENCE 178 AA; 20294 MW; F6A54D98A6B53211 CRC64;

```

lgment Scores:

```

Pred. No.: 0.000743 Length: 178
Score: 111.50 Matches: 49
Percent Similarity: 32.77% Conservative: 9
Best Local Similarity: 27.68% Mismatches: 60
Query Match: 14.41% Indels: 59
DB: 1 Gaps: 6

US-09-518-842-1 (1-420) x RELH_RABBIT (1-178)

QY 4 GCAGCGCTCTCCGGTCTCTATCGCAGCAATCTGGCTGCTGCTGAGCAA----- 54
Db 3 AlaLeuLeuPhe-----TyrLeuLeuGlyPheCysLeuGlnGlyValThrGly 20
QY 55 ----CTCCTTAGAGAAAGCTACGACAGAGCTGAGGGATGTGTCCTCCGATTTGGAATA 111
Db 21 ArgValThrTyrGluTrpMetMetGluAsnValLysIleCysArgAsnAaspPheValArg 40
QY 112 CACTTGCTGTCATATTGC-----CCCATGCTCGAGAAG 144
Db 41 ThrAlaIleGluValCysGlyHisValHisLeuGluArgGluSerProSerProGluAsn 60
QY 145 ACATTCACCAACCCAGGAGGGTGGCTGGGAATCTGGACGCTCCCAAGAAATGGTG 204
Db 61 ProPhe-----LeuSerSerGlyProAlaAlaGluThrVal 72
QY 205 TCACCTCCACACACACAGATGGACAGCCTTAGTAGACATCAGATTCATTCCTTAAT 264
Db 73 ProSerSerIleLysLysAaspAlaGluAsnAlaAsnThrMetLeuGluSerIleProAsn 92
QY 265 TTGTACACAGAGCTGAAGAACCACTGTCTGAAGGCGAGCCATCA----- 309
Db 93 LeuProGlnGluLeuThrAlaThrLeuPheGluLysGlnProSerLysLeuTyrLeuGln 112
QY 310 -----TTGAACAAAATAATA--- 324
Db 113 TyrLeuProThrLeuLysLysSerAsnValSerPheGluGluPheLysLysIleIleGln 132
QY 325 -----CTTCCCGCAAA 336
Db 133 AsnIleGlnArgGlyValGlnGlySerSerAlaSerGluSerAsnThrPheSerArgLys 152
QY 337 AAGAGAAGTGGACGTCACAGATTGATCCATTTCTGTGTGAAGTAATTTGT 387
Db 153 LysArgGlnPheSerGluSerLeuProGluGluCysCysLysTyrGlyCys 169

RESULT 11
RELH_PANTR
ID RELH_PANTR STANDARD; PRT; 166 AA.
AC P51454;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Prorelaxin H1 precursor (Fragment).
GN RNL1 OR RLX1.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=94238260; PubMed=8182365;
RA Evans B.A., Fu P., Tregear G.W.;
RT "Characterization of two relaxin genes in the chimpanzee."
RL J. Endocrinol. 140:385-392(1994).
CC -!- FUNCTION: RELAXIN IS AN OVARIAN HORMONE THAT ACTS WITH ESTROGEN TO
CC PRODUCE DILATION OF THE BIRTH CANAL IN MANY MAMMALS. MAY BE
CC INVOLVED IN REMODELING OF CONNECTIVE TISSUES DURING PREGNANCY,
CC PROMOTING GROWTH OF PUBIC LIGAMENTS AND RIPENING OF THE CERVIX.
CC -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN THE CORPUS LUTEUM OF PREGNANCY

```



```

2Y 1 ATGGCCAGCGCTGTCGGTCTATTCGACGAAATCGCTGCTGCTGACGCAACTCCTT 60
2b 1 MetLeuArgTrpPheLeuSerHisLeuLeuGlyValTrpLeuLeuLeuSerGlnLeuPro 20
2Y 61 AGAGAAAGCCTAGCAGCAGAG-----CTGAGGGGATGTGGTCCCGCATTTGGAAAA 111
2b 1 MetLeuArgTrpPheLeuSerHisLeuLeuGlyValTrpLeuLeuLeuSerGlnLeuPro 20
2Y 21 ArgGluLeuProAlaThrAspAspLysLysLeuLysAlaCysGlyArgAspTyrValArg 40
2b 21 ArgGluLeuProAlaThrAspAspLysLysLeuLysAlaCysGlyArgAspTyrValArg 40
2Y 112 CACTTGCTCATATTGCCCCATGCTGAGAGAGACATTACACACCCAGGAGG--- 168
2b 112 CACTTGCTCATATTGCCCCATGCTGAGAGAGACATTACACACCCAGGAGG--- 168
2Y 41 LeuGlnIleGluValCys-----GlySerIle 49
2b 41 LeuGlnIleGluValCys-----GlySerIle 49
2Y 169 ----TGCTGCTGGAATCTGGAGTCCCAAA----- 195
2b 169 ----TGCTGCTGGAATCTGGAGTCCCAAA----- 195
2Y 50 TrpTrpGlyArgLysAlaGlyGlnLeuArgGluArgArgGlnIleSerGluProLeuAla 69
2b 50 TrpTrpGlyArgLysAlaGlyGlnLeuArgGluArgArgGlnIleSerGluProLeuAla 69
2Y 196 GAATGGTGTCAACCTCCCAACAAAGAGTGGACAGCTTAGGTACGACATCAGAAATTC 255
2b 196 GAATGGTGTCAACCTCCCAACAAAGAGTGGACAGCTTAGGTACGACATCAGAAATTC 255
2Y 70 GluValValProSerSerIleLeuAsnAspProGluIleLeuSerLeuMetLeuGlnSer 89
2b 70 GluValValProSerSerIleLeuAsnAspProGluIleLeuSerLeuMetLeuGlnSer 89
2Y 256 ATTCCTAATTTGTCACGAGCTGAAGAAACCACTGCTGAGAGGGCAGCCATCATTC--- 312
2b 256 ATTCCTAATTTGTCACGAGCTGAAGAAACCACTGCTGAGAGGGCAGCCATCATTC--- 312
2Y 90 IleProGlyMetProGlnGluLeuArgIleAlaThrArgSerGlyLysGluLeuLeu 109
2b 90 IleProGlyMetProGlnGluLeuArgIleAlaThrArgSerGlyLysGluLeuLeu 109
2Y 313 -----AGAAAAATA 321
2b 313 -----AGAAAAATA 321
2Y 110 ArgGluLeuHisPheValLeuGluAspSerAsnLeuAsnLeuGluGluMetLysLysThr 129
2b 110 ArgGluLeuHisPheValLeuGluAspSerAsnLeuAsnLeuGluGluMetLysLysThr 129
2Y 322 ATACTTTCC----- 330
2b 322 ATACTTTCC----- 330
2Y 130 PheLeuAsnThrGlnPheGluAlaGluAspLysSerLeuSerLysLeuAspLysHisPro 149
2b 130 PheLeuAsnThrGlnPheGluAlaGluAspLysSerLeuSerLysLeuAspLysHisPro 149
2Y 331 CGCAAAAGAGAGTGGAGCTCACAGATTGATTCATTCCTGTTGGAAGTAATTTGT 387
2b 331 CGCAAAAGAGAGTGGAGCTCACAGATTGATTCATTCCTGTTGGAAGTAATTTGT 387
2Y 150 ArgLysLysArgAspAsnTyrIleLysMetSerAspLysCysAsnValGlyCys 168
2b 150 ArgLysLysArgAspAsnTyrIleLysMetSerAspLysCysAsnValGlyCys 168

```

RESULT 13

```

ELX_MESAU
D RELX_MESAU STANDARD; PRT; 177 AA.
C 064171;
T 01-NOV-1997 (Rel. 35, Created)
T 01-NOV-1997 (Rel. 35, Last sequence update)
T 16-OCT-2001 (Rel. 40, Last annotation update)
E Prorelaxin precursor.
N RLN.
S Mesocricetus auratus (Golden hamster).
C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
C Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
C Mesocricetus.
X NCBI_TaxID=10036;
N [1]
P SEQUENCE FROM N.A.
C TISSUE=Placenta;
X MEDLINE=96115021; PubMed=7492700;
A McCaslin R.B., Renegar R.H.;
T "Determination of the prorelaxin nucleotide sequence and expression
of prorelaxin messenger ribonucleic acid in the golden hamster.";
L Biol. Reprod. 53:454-461(1995).
C -1- FUNCTION: RELAXIN IS AN OVARIAN HORMONE THAT ACTS WITH ESTROGEN TO
PRODUCE DILATATION OF THE BIRTH CANAL IN MANY MAMMALS. IT BEARS
NATURE YOUNG, AND ALLOWS SEPARATION OF THE PELVIC BONES.
C -1- SUBUNIT: HETERO DIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
DISULFIDE BONDS.
C -1- CELLULAR LOCATION: Secreted.
C -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.

```

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

or send an email to license@isb-eib.ch.

```

CC EMBL; S79879; AAB35655.1; -
DR InterPro; IPR004825; Ins/IGF/relax.
DR Pfam; PF00049; Insulin; 1.
DR SMART; SM00078; IIGF; 1.
DR PROSITE; PS00262; INSULIN; 1.
KW Insulin family; Hormone; Signal.
FT SIGNAL 1 22 BY SIMILARITY
FT CHAIN 23 59 RELAXIN B CHAIN (PROBABLE).
FT PROPEP 64 149 CONNECTING PEPTIDE (PROBABLE).
FT CHAIN 154 177 RELAXIN A CHAIN (PROBABLE).
FT DISULFID 36 164 INTERCHAIN (BY SIMILARITY).
FT DISULFID 48 177 INTERCHAIN (BY SIMILARITY).
FT DISULFID 163 168 BY SIMILARITY.
SQ SEQUENCE 177 AA; 20007 MW; 6925562BD8C66CCD CRC64;

```

Alignment Scores:
Pred. No.: 0.0376 Length: 177
Score: 95.50 Matches: 33
Percent Similarity: 39.83% Conservative: 14
Best Local Similarity: 27.97% Mismatches: 40
Query Match: 12.34% Indels: 31
DB: 1 Gaps: 4

US-09-518-842-1 (1-420) x RELX_MESAU (1-177)

```

QY 1 ATGGCCAGCGCTGTCGGTCTATTCGACGAAATCGCTGCTGCTGACGCAACTCCTT 60
Db 1 MetSerCysLysPheValLeuGlnLeuLeuGlyPheTrpLeuLeuLeuSerGlnProCys 20
QY 61 AGAGAAAGCCTAGCAGCAGAG-----CTGAGGGGATGTGGTCCCGCATTT 105
Db 21 ArgAlaArgValThrLysGluTrpLeuAspLysValIleHisValCysGlyArgGluTyr 40
QY 106 GCAAAACACTTGTGTCATATTC-----CCCATGCCCT 138
Db 41 ValArgAlaIleLeuAspIleCysAlaAlaThrValGlyLeuGluAlaProProLeuArg 60
QY 139 GAGAGACATTCACCAACCCAGGAGGGTGGTCTGGAATCTGGACGTCCTCAAGAA 198
Db 61 ArgArgArgMetThr-----GluGlu 67
QY 199 ATGGTGTCAACCTCCCAACAAAGATGGACAGCTTAGGTACGACATCAGAAATTCATT 258
Db 68 AlaValSerSerPheIleLysGluAspAlaGluProPheAspThr-----Met 83
QY 259 CTAATTTGTCACGAGCTGAAGAAACCACTGCTGAGAGGGCAGCCATCATTTG 312
Db 84 ProAsnLeuSerGluLysProLysThrAlaLeuProGluGlyHisProSerLeu 101

```

RESULT 14

```

IBP6_HUMAN
ID IBP6_HUMAN STANDARD; PRT; 240 AA.
AC P24592; Q14492;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Insulin-like growth factor binding protein 6 precursor (IGFBP-6)
DE (IBP-6) (IGF-binding protein 6).
GN IGFBP6 OR IBP6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Osteosarcoma;
RX MEDLINE=91225006; PubMed=1709161;
RA Kiefer M.C., Maslarz F.R., Bauer D.M., Zapf J.;
RT "Identification and molecular cloning of two new 30-kDa insulin-like
growth factor binding proteins isolated from adult human serum.";
J. Biol. Chem. 266:9043-9049(1991).

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

M nucleic - protein search, using frame_plus_n2p model

un on: October 9, 2003, 12:12:26 ; Search time 31 Seconds

(without alignments)
2605.861 Million cell updates/sec

title: US-09-518-842-1

erfect score: 774

sequence: 1 ATGCCAGCCTGTTCCGTC.....CAGTAAATTATTCATACATG 420

coring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

earched: 283308 seqs, 96168682 residues

total number of hits satisfying chosen parameters: 566616

inimum DB seq length: 0

aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

ommand line parameters:

MODEL=frame+n2p.model -DEV=xlp

Q=/cgn21/USPTO-spool/US09518842/runat_09102003_111037_13078/app.query.fasta_1.583

DB=PIR_76 -QMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0

UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45

DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL

OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000

USR=US09518842 @CGN 1 1 62 @runat_09102003_111037_13078 -NCPU=6 -ICPU=3

NO MMAP -LARGEOUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG

DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

atabase :

PIR_76:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Query Match	Length	ID	Description
1	180	23.3	182	1	RXP
2	151	19.5	185	2	S48082
3	123.5	16.0	185	1	A60982
4	122.5	15.8	185	2	A34936
5	120	15.5	186	1	RXRT
6	119.5	15.4	185	1	A44559
7	113.5	14.7	166	2	S42786
8	111.5	14.4	178	2	A49014
9	108.5	14.0	166	2	S42783
10	103.5	13.4	162	2	T16753
11	101	13.0	560	2	T16833
12	94	12.6	754	2	A20614
13	83	10.7	321	2	E83259
14	82.5	10.7	1626	2	T09271

C	15	82	11.0	335	1	T02809	probable peptidylp
	16	80.5	10.4	457	2	AG0433	proteinase (EC 3.4
	17	80	10.3	240	2	A39842	insulin-like growt
C	18	77.5	10.4	1217	2	T25894	hypothetical prote
	19	77	9.9	1490	2	S72351	nonstructural poly
	20	76.5	9.9	160	2	A49194	relaxin - guinea p
C	21	76.5	10.3	402	2	A84581	probable disease f
	22	76	9.8	188	2	JC6547	high sulfur protei
	23	75.5	9.8	391	2	S60672	replication protei
	24	75	9.7	224	1	B40850	secretory protein
C	25	75	10.1	993	2	S46779	26S proteasome reg
	26	75	10.1	1711	1	A55148	protein-tyrosine-p
C	27	74.5	10.0	564	2	T38291	GATA-type transcri
	28	74.5	10.0	564	2	T43298	transcription fact
	29	74.5	9.6	1531	2	H71468	probable outer mem
	30	73.5	9.5	273	2	G89979	NAD synthetase, pr
C	31	73.5	9.9	296	1	OPNB7	peroxidase (EC 1.1
	32	73.5	9.5	453	2	I50674	retinoic acid rece
C	33	73	9.8	315	2	T10252	peroxidase (EC 1.1
	34	73	9.4	1435	2	S69632	regulatory protein
C	35	73	9.8	1744	1	C4HU	complement C4A pre
	36	73	9.4	2493	2	S26372	nonstructural poly
	37	73	9.4	2493	2	S72349	nonstructural poly
	38	72.5	9.4	292	2	AB0825	probable membrane
C	39	72.5	9.4	1259	2	T16038	hypothetical prote
	40	72	9.7	403	2	T20070	hypothetical prote
C	41	72	9.7	443	2	AI3525	4-aminobutyrate tr
	42	72	9.3	479	2	B87092	GTPI/Obg-family Gr
	43	72	9.3	483	2	AD3618	succinate-semialde
	44	72	9.3	752	2	T50450	hypothetical prote
	45	72	9.3	836	1	JDVLD	DNA-directed DNA p

ALIGNMENTS

RESULT 1

RXP

relaxin precursor - pig

C:Species: Sus scrofa domestica (domestic pig)

C>Date: 02-Apr-1982 #sequence revision 02-Apr-1982 #text change 18-Jun-1999

C:Accession: A90934; A93187; A90205; A90196; S32313; S32312; A29796; A01615

R:Hailey, J.; Hudson, P.; Scanlon, D.; John, M.; Cronk, M.; Shine, J.; Tregear, G.; Nial

DNA 1, 155-162, 1982

A:Title: Porcine relaxin: molecular cloning and cDNA structure.

A:Reference number: A90934; MUID:83157118; PMID:6897721

A:Accession: A90934

A:Molecule type: DNA

A:Residues: 1-182 <HAL>

A:CROSS-references: GB:K01088; NID:gl64634; PIDN:AAA31114.1; PID:gl64635

R:James, R.; Niall, H.; Kwok, S.; Bryant-Greenwood, G.

Nature 267, 544-546, 1977

A:Title: Primary structure of porcine relaxin: homology with insulin and related growth

A:Reference number: A93187; MUID:77213067; PMID:876374

A:Accession: A93187

A:Molecule type: protein

A:Residues: 25-50, 'TWGR', 161-182 <JAM>

R:Schwabe, C.; McDonald, J.K.; Steinetz, B.G.

Biochem. Biophys. Res. Commun. 75, 503-510, 1977

A:Title: Primary structure of the B-chain of porcine relaxin.

A:Reference number: A90205; MUID:77157271; PMID:851452

A:Accession: A90205

A:Molecule type: protein

A:Residues: 25-47, 'VW', 50 <SCH1>

R:Schwabe, C.; McDonald, J.K.

Biochem. Biophys. Res. Commun. 74, 1501-1504, 1977

A:Title: Demonstration of a pyroglutamate residue at the N terminus of the B-chain of po

A:Reference number: A90201; MUID:77134136; PMID:843375

A:Contents: annotation; pyroglutamate carboxylic acid

R:Schwabe, C.; McDonald, J.K.; Steinetz, B.G.

Biochem. Biophys. Res. Commun. 70, 397-405, 1976

A:Title: Primary structure of the A chain of porcine relaxin.

A:Reference number: A90196; MUID:76231539; PMID:938497

A:Accession: A90196

Df

111 LeuGlnSerAlaSerLysAspSerAsnLeuAenPheGluGluPheLysLeilelle 130

325 CTT----- 327

131 LeuAsnArgGlnAsnGluAlaGluAspLysSerLeuLeuGluLeuLysAsnLeuGlyLeu 150

328 -----TCCGGCAAAAGAGAAGTGACGTGCACAGATTTCATCCTGTTGTGA 378

151 AspLysHisSerArgLysIArgLeuPheArgMetThrLeuSerGluLysCysCysGln 170

379 GTAATTGGTGACCATGGAACCTTCAGTTAAATTATGT 414

171 ValGlyCysIleargLysAspIleAlaArgLeuCys 182

RESULT 2

S48082

relaxin precursor - mouse

C:Species: Mus musculus (house mouse)

C>Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 16-Jul-1999

C:Accession: S48082; PC2067; FN0626

R:Ivans, B.A.; John, M.; Fowler, K.J.; Summers, R.J.; Cronk, M.; Shine, J.; Tregear, G.

J. Mol. Endocrinol. 10, 15-23, 1993

A>Title: The mouse relaxin gene: nucleotide sequence and expression.

A:Reference number: S48082; MUID:93199663; PMID:8452637

A:Accession: S48082

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-185 <BUE>

A:Cross-references: EMBL:Z27088; NID:G414780; PIDN:CAA81611.1; PID:G414781

R:Buellesbach, E.E.; Schwabe, C.

Biochem. Biophys. Res. Commun. 196, 311-319, 1993

A>Title: Mouse relaxin: synthesis and biological activity of the first relaxin with an

A:Reference number: PN0626; MUID:94030011; PMID:8216305

A:Accession: FC2067

A:Molecule type: protein

A:Residues: 23-57 <BUE>

A:Accession: PN0626

A:Molecule type: protein

A:Residues: 161-185 <BU2>

A>Note: proteins with and without 184-Tyr were synthesized, their biological activities

C:Superfamily: insulin

F:23-57,161-185/Product: relaxin #status experimental <MAT>

F:23-57/Domain: chain B #status experimental <CHB>

F:161-185/Domain: chain A #status experimental <CHA>

F:36-171,48-185,170-175/Disulfide bonds: #status experimental

Alignment Scores:

Pred. No.:	Score:	Length:
1	8.34e-08	185
Percent Similarity:	151.00	Matches: 50
Best Local Similarity:	37.70%	Conservative: 19
Query Match:	27.32%	Mismatches: 68
DB:	19.51%	Indels: 46
	2	Gaps: 3

US-09-518-842-1 (1-420) x S48082 (1-185)

QY 1 ATGCCGAGCTGTTCGGTCTTATCGCCAGCAATCTGGCTGTCTGTGACCCAACACTCTT 60

Df 1 MetSerArgPheLeuLeuGlnLeuLeuGlyPheThrLeuLeuSerGlnProCys 20

QY 61 AGAGAAAGCCTACACAGAG-----CTGAGGGGATCTGTCCTCCGATTT 105

Df 21 ArgThrArgValSerGluGluTrpMetAspGlyPheIleArgMetCysGlyArgGluTyr 40

QY 106 GGAAAAACATTGCTGTCATATTCCCOCATTCCTGAGAAGACATTCACCACCCCCAGGA 165

Df 41 AlaArgGluLeuIleLysIleCysGlyAlaSerValGlyArgLeuAlaLeuSerGlnGlu 60

QY 166 GGGTGGCTGTGGAATCTGGACCTCCCAAGAAATGGTGTCAACCTCCACACAAGAT 225

Df 61 GluProAlaLeuLeuAlaArgGlnAlaThrGluValValProSerPheIleAsnLysAsp 80

```

226 GGACAAGCCTTAGTACGACATCAGAAATTCATTCTTAATTTGTCCACGAGCTGAAGAAA 285
    :
81 AlaGluProPheAspThrThrLeuLysCysLeuProHisLeuSerGluGluLeuLysAla 100
    :
286 CCACTGCTCTGAAGGGCGCCATTCATTTGAAGAAATA 327
    :
101 ValLeuSerGluAlaGlnAlaSerLeuProGluLeuGlnHisAlaProValLeuSerAsp 120
    :
327 -----
121 SerValValSerLeuGluGlyPheLysThrLeuHisAspArgLeuGlyGluAlaGlu 140
328 -----TCCGCAAAAAGAGA 342
    :
141 AspGlySerProProGlyLeuLysTyrLeuGlnSerAspThrHisSerArgLysLysArg 160
343 AGTGGAGCTCACAGATTGATCCATTCCTTTGGAAGTAATTTGTGACGATGGAACCTTCA 402
    :
161 GluSerGlyGlyLeuMetSerGlnGlnCysCysHisValGlyCysSerArgArgSerIle 180
403 GTTAATAATTA 411
    :
181 AlaLysLeu 183

RESULT 3
60982
;elaxin 2 precursor [validated] - human
;Alternate names: preprorelaxin 2
;Species: Homo sapiens (man)
;Date: 03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change 08-Dec-2000
;Accession: A05092; A60982
;Hudson, P.; John, M.; Crawford, R.; Hatlamabidis, J.; Scanlon, D.; Gorman, J.; Tregear, J.
;MBO J. 3, 2333-2339, 1984
;Title: Relaxin gene expression in human ovaries and the predicted structure of a human
;Reference number: A05092; MUID:65051298; PMID:6548702
;Accession: A05092
;Molecule type: mRNA
;Residues: 1-185 <HUD>
;Cross-references: GB:X00948; NID:G35926; PID:CAA25460.1; PID:G35927
;Stults, J.T.; Borell, J.H.; Canova-Davis, E.; Ling, V.T.; Laramee, G.R.; Winslow, J.W.
;J. Biol. Chem. 265, 655-664, 1990
;Title: Structural characterization by mass spectrometry of native and recombinant human
;Reference number: A60982; MUID:91167739; PMID:2076464
;Accession: A60982
;Molecule type: protein
;Residues: 25-53;162-185 <STU>
;Genetics:
;Gene: GDB:RLN2
;Cross-references: GDB:119553; OMIM:179740
;Map position: 9pter-9q12
;Superfamily: insulin
;Keywords: ovary; pyroglutamic acid
;1-24/Domain: signal sequence #status predicted <SIG>
;25-53/Domain: relaxin 2 chain B #status experimental <BCH>
;25-53,162-185/Product: relaxin 2 #status experimental <MAT>
;58-157/Domain: relaxin 2 connecting C peptide #status predicted <CPDP>
;162-185/Domain: relaxin 2 chain A #status experimental <ACH>
;35-172,47-185,171-176/Disulfide bonds: #status experimental
;162/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimental

Alignment Scores:
red. NO.: 7,45e-05 Length: 185
core: 123.50 Matches: 53
Percent Similarity: 35.35% Conservative: 17
Best Local Similarity: 26.77% Mismatches: 55
Query Match: 15.96% Indels: 73
B: 1 Gaps: 7

S-09-518-842-1 (1-420) x A60982 (1-185)
Y 1 ATGGCCAGCCTGTTCCGGTCTCTATCTGCCAGCAATCTGGCTGCTGCTGAGCCAACTCCTT 60
b 1 MetProArgLeuPhePheHisLeuLeuGlyValCysLeuLeuLeuAenGlnPheSer 20

```

```

QY 61 -----AGAGAAAGCCTA 72
Db 21 ArgAlaValAlaAspSerTrpMetGluGluValIleLysLeuCysGlyArgGluLeuVal 40
QY 73 GCAGCAGAGCTCAGGGGATGTGGT---CCCCGATTGGAAAAACACTTGTGTGCATATTGC 129
Db 41 ArgAlaGlnIleAlaAlaCysGlyMetSerThrTrpSerLysArgSerLeuSerGlnGlu 60
QY 130 CCCATGCTCTGAGAGACATCTCACACACCCAGAGGGTGGCTCTCGGAATCTCGAGCT 189
Db 61 AspAlaProGln-----ThrPro-----Arg 67
QY 190 CCC---AAGAAATCGTGTCAACCTCCACACAAAGATGGACAGCCTTAGGTACGACA 246
Db 68 ProValAlaGluIleValProSerPheIleAsnLysAspThrGluThrIleAsnMetMet 87
QY 247 TCAGAATTCATTCCTAAATTTGTCCACAGAGCTGGAAGAAACCACTGTCTGAAGGGCAGCA 306
Db 88 SerGluPheValAlaAsnLeuProGlnGluLeuLysLeuThrLeuSerGluMetGlnPro 107
QY 307 TCATTG----- 312
Db 108 AlaLeuProGlnLeuGlnHisValProValLeuLysAspSerSerLeuLeuPheGlu 127
QY 313 -----AAGAAATTAATATT----- 327
Db 128 GluPheLysLysLeuIleArgAsnArgGlnSerGluAlaAlaAspSerSerProSerGlu 147
QY 328 -----TCCGCAAAAAGAGAAGTGGACGTCACAGATTT 360
Db 148 LeuLysTyrLeuGlyLeuAspThrHisSerArgLysLysArgGlnLeuTyrSerAlaLeu 167
QY 361 GATCATTCTGTTGGAAGTAATTTGTCCAGCATGGAACCTCAGTTAAATATGT 414
Db 168 AlaAsnLysCysCysHisValGlyCysThrLysArgSerLeuAlaArgPheCys 185

RESULT 4
A34936
relaxin precursor - rhesus macaque
C:Species: Macaca mulatta (rhesus macaque)
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 12-Jun-1998
R:Crawford, R.J.; Hammond, V.E.; Roche, P.J.; Johnston, P.D.; Tregear, G.W.
J. Mol. Endocrinol. 3, 169-174, 1989
A:Title: Structure of rhesus monkey relaxin predicted by analysis of the single-copy rh
A:Reference number: A34936; MUID:90073957; PMID:2590381
A:Accession: A34936
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-185 <CRA>
C:Superfamily: insulin
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-185/Product: relaxin #status predicted <MAT>

Alignment Scores:
Pred. No.: 9,53e-05 Length: 185
Score: 122.50 Matches: 51
Percent Similarity: 34.34% Conservative: 17
Best Local Similarity: 25.76% Mismatches: 57
Query Match: 15.83% Indels: 73
DB: 2 Gaps: 6

US-09-518-842-1 (1-420) x A34936 (1-185)
QY 1 ATGGCCAGCCTGTTCCGGTCTCTATCTGCCAGCAATCTGGCTGCTGCTGAGCCAACTCCTT 60
Db 1 MetProArgLeuPhePheHisLeuLeuGlyValCysLeuLeuLeuAenGlnPheSer 20
QY 61 -----AGAGAAAGCCTA 72
Db 21 ArgAlaValAlaAlaLysTrpMetAspValIleLysAlaCysGlyArgGluLeuVal 40

```

```

QY 73 GCAGCAGACTGAGGGATGTGT---CCCCGATTGGAAACACTTCTGCTCATATTGC 129
Db |||:||||| |||:||||| |||:||||| |||:|||||
41 ArgAlaGlnIleAlaIleCysGlySerThrLeuGlyLysArgSerLeuAsnGlnGlu 60
QY 130 CCATGCTGAGAGACATTCACACACCCAGGAGGCTGCTGGAATCTGGACGT 189
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
61 AspAlaProLeuLysPro-----Arg 67
QY 190 CCC---AAAGAAATGTGTCACTCCCAACAAAGATGACAGCCTTAGGTACGACA 246
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
68 ProAlaAlaGluIleValProSerLeuLeuAsnGlnAspThrGluThrIleAsnMetMet 87
QY 247 TCAGAAATTCCTTAATTTGCACAGCTGAAGAAACCACTGTCTGAAGGCGACCA 306
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
88 SerGluPheValAlaAsnLeuProGlnGluLeuLysLeuThrLeuSerGluArgGlnPro 107
QY 307 TCATTG----- 312
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
108 AlaLeuSerGluLeuGlnGlnHisValProValLeuLysAspSerAsnLeuSerPheGlu 127
QY 313 -----AAGAAATTAATCTT----- 327
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
128 GluPheLysLysIleLeuArgLysArgGlnSerGluAlaThrAspSerProSerGlu 147
QY 328 -----TCCCGAAAGAGAGAGTGGACGTCACAGATT 360
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
148 LeuArgSerLeuGlyLeuAspThrHisSerArgLysArgGlnLeuTyMetThrLeu 167
QY 361 GATCCATTCTGTTGCAAGTAATTTGTACGACGAACTTCAGTTAAATTTATGT 414
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
168 SerAsnLysCysCysHisIleGlyCysThrLysLysSerLeuAlaLysPheCys 185

RESULT 5
RXRT
relaxin precursor - rat
;Species: Rattus norvegicus (Norway rat)
;Date: 18-Dec-1981 #sequence_revision 18-Dec-1981 #text_change 24-Sep-1999
;Accession: A01614
;Hudson, P.; Haley, J.; Cronk, M.; Shine, J.; Niall, H.
ature 291, 127-131, 1981
;Title: Molecular cloning and characterization of cDNA sequences coding for rat relaxin
;Reference number: A01614; PMID:81197624; PMID:7231533
;Accession: A01614
;Molecule type: mRNA
;Residues: 1-186 <HUD>
;Cross-references: GB:J00780; GB:M25468; NID:g206606; PIDN:AAA42029.1; PID:g206607
;Superfamily: insulin
;Keywords: hormone; ovary; pyroglutamic acid
;1-22/Domain: signal sequence #status predicted <SIG>
;23-57/Domain: relaxin chain B #status predicted <RXB>
;23-57,163-186/Product: relaxin #status predicted <MAT>
;58-162/Domain: relaxin connecting C peptide #status predicted <RXC>
;163-186/Domain: relaxin chain A #status predicted <RXA>
;16-173,48-186,172-177/Disulfide bonds: #status predicted
;163/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted

Alignment Scores:
red. No.: 0.000177 Length: 186
core: 120.00 Matches: 46
percent Similarity: 33.16% Conservatives: 18
est Local Similarity: 23.83% Mismatches: 67
very Match: 15.50% Indels: 62
B: 1 Gaps: 5

S-09-518-842-1 (1-420) x RXRT (1-186)
Y 1 ATGCCAGCTGTTCGGTCTTATTCGCCAGCAATCTGCTGCTCTGCTGAGCAACTCTT 60
b |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
1 MetSerArgLeuLeuLeuLeuLeuGlyPheTrpLeuPheLeuSerGlnProCys 20
Y 61 AGAGAAAGCTACGACGAG-----CTGAGGGGATGTGTCTCCCGGATT 105
b |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
21 ArgAlaArgValSerGluGluTrpMetAspGlnValIleGlnValCysGlyArgGlyTyr 40

```

```

QY 106 GGAAACACTTGTCTCATATTGCCCATGCTGAGAAGACATTCACCAACCCACGGA 165
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
41 AlaArgAlaTrpIleGluValCys-----GlyAlaSerValGly 53
QY 166 GGGTGGCTGTGGAAATCTGGACGTCCTCC-----AAAGAAATG 201
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
54 ArgLeuAlaLeuSerGlnGluGluProAlaProLeuAlaArgGlnAlaThrAlaGluVal 73
QY 202 GTGTCAACTCCCAACAAAGATGGAACCTTAGTCAGCATCAGAAATTCATTCCT 261
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
74 ValProSerPheIleAsnLysAlaGluProPheAspMetThrLeuLysCysLeuPro 93
QY 262 AATTTGTCCAGAGCTCAAGAAACCACTGTCTGAAGGCGAG----- 303
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
94 AsnLeuSerGluGluArgLysAlaAlaLeuSerGluGlyArgAlaProPheProGluLeu 113
QY 303 ----- 303
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
114 GlnGlnHisAlaProAlaLeuSerAspSerValValSerLeuGluGlyPheLysThr 133
QY 304 -----CCATCATTGAAGAAATA----- 321
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
134 PheHisAsnGlnLeuGlyGluAlaGluAspGlyGlyProGluLeuLysTyLeuGly 153
QY 322 -----ATACTTCCCGCAAGAGAGAGTGGAGCTCACAGATTGATCCATTCCTTGT 375
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
154 SerAspAlaGlnSerArgLysArgGlnSerGlyAlaLeuLeuSerGluGlnCysCys 173
QY 376 GAAGTAATTTGTGACGATGCAACTTCAGTTAAATTTATGT 414
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
174 HisIleGlyCysThrArgArgSerIleAlaLysLeuCys 186

RESULT 6
A44559
N;Alternate names: preprorelaxin 1
C;Species: Homo sapiens (man)
C;Date: 03-Feb-1994 #sequence_revision 07-Oct-1994 #text_change 18-Jun-1999
C;Accession: B05092, A44559
R;Hudson, P.; John, M.; Crawford, R.; Haralambidis, J.; Scanlon, D.; Gorman, J.; Tregear,
EMBO J. 3, 2333-2339, 1984
A;Title: Relaxin gene expression in human ovaries and the predicted structure of a huma
A;Reference number: A05092; PMID:85051298; PMID:6548702
A;Accession: B05092
A;Molecule type: DNA
A;Residues: 1-185 <HUI>
A;Cross-references: GB:X00949; NID:g35932; PIDN:CAA25461.1; PID:g35933
R;Hudson, P.; Haley, J.; John, M.; Cronk, M.; Crawford, R.; Haralambidis, J.; Tregear,
Nature 301, 628-631, 1983
A;Title: Structure of a genomic clone encoding biologically active human relaxin.
A;Reference number: A44559; PMID:83141755; PMID:6298628
A;Accession: A44559
A;Molecule type: DNA
A;Residues: 1-185 <HU2>
A;Cross-references: GB:X00949; NID:g35932; PIDN:CAA25461.1; PID:g35933
C;Comment: Relaxin is an ovarian hormone that acts with estrogen to produce dilation of
disulfide bonds.
C;Genetics:
A;Gene: GDB:RLN1
A;Cross-references: GDB:119552; OMIM:179730
A;Map position: 9pter-9q12
C;Superfamily: insulin
C;Keywords: hormone; ovary
F;1-25/Domain: signal sequence #status predicted <SIG>
F;26-57,162-185/Domain: relaxin 1 chain B #status predicted <BCH>
F;58-158/Domain: relaxin 1 connecting C peptide #status predicted <MAT>
F;162-185/Domain: relaxin 1 chain A #status predicted <ACH>
F;35-172,47-185,171-176/Disulfide bonds: #status predicted

Alignment Scores:
Pred. No.: 0.0002 Length: 185

```

DB:	2	Gaps:	6
US-09-518-842-1 (1-420) x S42786 (1-166)			
QY	61	AGAGAAAGCCTAGCAGCAGAGCTGAGGGATGTGGT---CCCGATTGTTGAAACACTTG 117	
Db	18	ArgGluLeuValArgAlaGlnIleAlaIleCysGlySerThrTrpSerLysArgSer 37	
QY	118	CTGTCTATTTGCCCATGCTGTGAGAGACATTCCACCACCCACCCAGAGGGTGGCTGCTG 177	
Db	38	LeuSerGlnGluAspAlaProGln-----ThrPro----- 47	
QY	178	GAATCTGGAGCTCC---AAAGAAATGGTGTCAACCTCCAAACAACAAGATGCAAGCC 234	
Db	48	-----ArgProValAlaGluIleValProSerPheIleAsnLysAspThrGluThr 64	
QY	235	TTAGTAGCAGCATCAGAAATTCATTCTTAATTTGTCCACGAGCTGAAGAACCACTGTCT 294	
Db	65	IleAsnMetMetSerGluPheValAlaAsnLeuProGlnGluLeuLysLeuThrLeuSer 84	
QY	295	GAAGGGCAGCATCATG----- 312	
Db	85	GluMetGlnProAlaLeuProGlnLeuGlnTyrValProValLeuLysAspSerSer 104	
QY	313	-----AAGAAATAATACTT----- 327	
Db	105	LeuLeuPheGluGluPheLysLysLeuIleArgAsnArgGlnSerGluAlaAlaAspSer 124	
QY	328	-----TCCGCGCAAAAAGAGAGAGTGGGA 348	
Db	125	SerProSerGluLeuLysTyrLeuGlyLeuAspThrHisSerArgLysLysArgGlnLeu 144	
QY	349	CGTCACAGATTTGATCCATTCTGTGTGAAGTAATTTGTGACCATCGAACTTCAGTTAAA 408	
Db	145	TyrSerAlaLeuAlaAsnLysCysCysHisValGlyCysThrLysArgSerLeuAlaArg 164	
QY	409	TTATGT 414	
Db	165	PheCys 166	
RESULT 8			
A49014			
20K protein - rabbit			
C:Species: Oryctolagus cuniculus (domestic rabbit)			
C>Date: 19-Dec-1993 #sequence_revision 19-Nov-1994 #text_change 16-Jul-1999			
C:Accession: A49014			
R:Jettett, A.M.; Bernacki, S.H.; Floyd, E.E.; Saunders, N.A.; Pieniazek, J.; Lot			
Cell Growth Differ. 3, 549-556, 1992			
A:Title: Expression of a preprorelaxin-like gene during squamous differentiation			
A:Reference number: A49014; MUID:93002619; PMID:1339318			
A:Accession: A49014			
A>Status: preliminary			
A:Molecule type: nucleic acid			
A:Residues: 1-178 <JET>			
A:Cross-references: GB:S45940; NID:9257388; PIDN:AAB23648.1; PID:9257389			
A:Experimental source: tracheobronchial epithelial cells			
A:Note: sequence extracted from NCBI backbone (NCBIN:115816, NCBI:P:115821)			
C:Superfamily: insulin			
Alignment Scores:			
Pred. No.:	0.00144	Length:	178
Score:	111.50	Matches:	49
Percent Similarity:	32.77%	Conservative:	9
Best Local Similarity:	27.68%	Mismatches:	60
Query Match:	14.41%	Indels:	59
DB:	2	Gaps:	6
US-09-518-842-1 (1-420) x A49014 (1-178)			
QY	4	GCCAGCCTGTTCCGGTCTCTATCTCCAGCAATCTGGCTGTGTGAGCAAC----- 54	
Db	3	AlaLeuLeuPhe-----TyrLeuLeuGlyPheCysLeuLeuGlnGlyGlnValThrGly 20	

```

QY 55 ---CTCCTTAGAGAGCCTAGCAGCAGCTGAGGGATGTGTCCTCCGATTTGGAA 111
Db 21 ArgValThrTyrGluTrpMetMetGluAsnValIleCysArgAsnAspPheValArg 40
QY 112 CACTTGCTGCATATTCG- 144
Db 41 ThrAlaIleGluValCysGlyHisValHisLeuGluArgGluSerProGluAsn 60
QY 145 ACATTACACACCCAGAGGGTGGCTGTGGAATCTGGACCTCCCAAGAAATGGTG 204
Db 61 ProPhe- 72
QY 205 TCAACCTCCACACACAGAGAGCAACCTTGTAGTACGACATCAGAAATTCATCTCTAAT 264
Db 73 ProSerSerIleLeuLysLysAspAlaGluAsnAlaAsnThrMetLeuGluSerIleProAsn 92
QY 265 TTGTCCACAGAGCTGAGAAACCACTGTCTGAAGGGCAGCCATCA- 309
Db 93 LeuProGlnGluLeuThrAlaThrLeuPheGluLysGlnProSerLysLeuTyrLeuGln 112
QY 310 -----TTGAAGAAATAAATA- 324
Db 113 TyrLeuProThrLeuLysLysSerAsnValSerPheGluGluPheLysLysIleLeuGln 132
QY 325 -----CTTCCCGCAAA 336
Db 133 AsnIleGlnArgGlyValGlnGlySerSerAlaSerGluSerAsnThrPheSerArgLys 152
QY 337 AAGAGAGTGGAGCTCACAGATTTGATCCATCTGTTGTGAAGTAATTTGT 387
Db 153 LysArgGlnPheSerGluSerLeuProGluGluCysCysLysTyrGlyCys 169

RESULT 9
42783
relaxin 1 precursor - chimpanzee (fragment)
;Species: Pan troglodytes (chimpanzee)
;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Jul-1999
;Accession: S42783
;Evans, B.B.
submitted to the EMBL Data Library, November 1993
;Reference number: S42776
;Accession: S42783
;Molecule type: mRNA
;Residues: 1-166 <EVA>
;Cross-references: EMBL:Z27225; NID:g415996; PIDN:CAA81739.1; PID:g415997
;Genetics:
;Gene: rlx1
;Superfamily: insulin
;Keywords: disulfide bond; hormone
;1-5/Domain: signal sequence (fragment) #status predicted <SIG>
;16-166/Product: relaxin 1 #status predicted <MAT>

Alignment Scores:
red. No.: 0.00299 Length: 166
core: 108.50 Matches: 43
Percent Similarity: 37.65% Conservative: 18
Best Local Similarity: 26.54% Mismatches: 44
Query Match: 14.02% Indels: 57
B: 2 Gaps: 7

S-09-518-842-1 (1-420) x S42783 (1-166)

QY 61 AGAGAAAGCCTAGCAGCAGCTGAGGGATGTGCT---CCCGATTTGAAACACTTG 117
b 18 ArgGluLeuValArgAlaGlnIleAlaIleCysGlyMetSerThrTrpSerLysArgSer 37
Y 118 CTGTCATATTCCTCCATCGCTGAGAGACATTCACCCACCAGGGGTGCTGCTG 177
b 38 LeuSerGlnGluAspAlaProGln-----ThrPro----- 47
Y 178 GAATCTGAGAGCTGCC- --AAGAAATGGTGTCACCTCCACACACAAAGATGGACAGCC 234
b 48 -----ArgProValAlaGluIleValProSerPheIleAsnLysAspThrGluThr 64

```

```

QY 235 TTAGTAGACATCATGAATTCATTCCTTAATTTGTACACAGAGCTGAAGAAACCACTGCT 294
Db 65 IleIleIleMetLeuGluPheIleAlaAsnLeuProGluLeuLysAlaAlaLeuSer 84
QY 295 GAAGGGCCATCATTCG- 312
Db 85 GluArgGlnProSerLeuProGluProGlnGlnTyrValProAlaLeuLysAspSerAsn 104
QY 313 -----AAGAAATAATATCTTCCCGCAAAAGAGAGT- 345
Db 105 LeuSerPheGluGluPheLysLysLeuIleArgAsnArgGlnSerGluAlaAlaAspSer 124
QY 346 -----GCACGTACACAGATTTGATCCA 366
Db 125 AsnProSerGluLeuLysTyrLeuGlyLeuAspThrHisSerGlnLysLysArgGlnPro 144
QY 367 TTC-----TGTGTGAAGTAAATTTGTGACGATGGAACTTCAGTTAAA 408
Db 145 TyrValAlaLeuPheGluLysCysLeuIleGlyCysThrLysArgSerLeuAlaAsn 164

RESULT 10
147053
relaxin B,C and A chains - horse (fragment)
;Species: Equus sp.
;Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 16-Jul-1999
;Accession: I47053
;Klonisch, T.; Ryan, P.L.; Yamashiro, S.; Porter, D.G.
Biol. Reprod. 52, 1307-1315, 1995
A:Title: Partial complementary deoxyribonucleic acid cloning of equine relaxin messenger
A:Reference number: I47053; MUID:95359320; PMID:7543295
A:Accession: I47053
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-143 <KLO>
A:Cross-references: GB:S78800; NID:gl042059; PIDN:AAB35036.1; PID:gl042060
C:Superfamily: insulin

Alignment Scores:
Pred. No.: 0.0101 Length: 143
Score: 103.50 Matches: 30
Percent Similarity: 51.76% Conservative: 14
Best Local Similarity: 35.29% Mismatches: 34
Query Match: 13.37% Indels: 7
DB: 2 Gaps: 3

US-09-518-842-1 (1-420) x I47053 (1-143)

QY 82 CTGAGGGATGTGTCCTCCGATTTGGAAACACTTGCTGTCATATTCGCCCATGCT- 138
Db 1 IleLysAlaCysGlyArgGluLeuAlaArgLeuArgIleGluIleCysGlySerLeuSer 20
QY 139 ---GAGAAGACATTC- ----ACCACACCCAGAGGGTGGCTGCTGGAATCTGGAGCT 189
Db 21 TrpLysLysThrValLeuArgLeuGluProGly-----LeuGluValGlyGln 37
QY 190 CCAAGAAATGGTGTCACCTCCACCAACAAAGATGGACAGCTTAGGTACGACATCA 249
Db 38 ProValGluIleValSerSerSerLysAspAlaGluAlaLeuAsnThrLysLeu 57
QY 250 GAATTCATTCCTAATTTGTCCAGAGCTGAGAAACACCTGCTCTGAAGGCGCCATCA 309
Db 58 GlyLeuAsnSerAsnLeuProLysGlnLysAlaThrLeuSerGluArgGlnProSer 77
QY 310 TTGAAGAAATAAATA 324
Db 78 TrpArgGluLeuLeu 82

RESULT 11

```

6833
hypothetical protein T07H6.5 - Caenorhabditis elegans
Species: Caenorhabditis elegans
Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Mar-2000
Accession: T16833
Geisel, C.
Submitted to the EMBL Data Library, April 1996
Description: The sequence of C. elegans cosmid T07H6.
Reference number: Z18586
Accession: T16833
Status: preliminary; translated from GB/EMBL/DBJ
Molecule type: DNA
Residues: 1-560 <GET>
Cross-references: ENBL:U53344; NID:g1255886; PID:g1255889; PIDN:AAA96225.1; GSPDB:GN00
Experimental source: strain Bristol N2; clone T07H6
Genetics:
Map position: X
Introns: 14/1; 75/1; 102/3; 128/1; 186/1; 272/2; 326/1; 361/1; 422/1; 475/1; 527/1
Superfamily: C4b-binding protein alpha chain; complement factor H repeat homology

Alignment Scores:

	Length:	Score:
0.0218	560	
101.00	43	
Conservative:	16	
Mismatch:	38	
Indels:	56	
Gaps:	9	

3-09-518-842-1 (1-420) x T16833 (1-560)

94	GTCCTCCCATTTGGCAAAACACTTGCTGTCATATTGCCCATCGCTGAGAACATTCACC	153	
414	GlyProPromeMet-----LysCysLeuProSerTrpCysGluHisProSerLysThrTyrGly	432	
154	ACCACCCCAGAGGGGTGGCTGCTGCTG	177	
433	ThrLeuProGlyGlyGlnIleLeuLeuGluGlyIleLeuGlyAlaTyrGluPheGlnSer	452	
178	-----GAATCTGGACGCTCCCAAGAAATGGTGTCAACCTCCAACAACAAA	222	
453	TyrIleGlnLysValGluGluGlyArg-----AlaIleSerPheGlnCysGlyLys	469	
223	GATGGACAAAGCCTTAGGT-----ACGACATCAGAAATTCATTCCTTAAT	264	
470	GlyAenTyrLeuIleGlyProProLysAlaThrCysValAenGlyGluTyrMetProLys	489	
265	TGTGCACACAGACTGAACAAACCACTGCTCAAGGCGCACCATCATTTGAAG---AAAAATA	321	
490	ValSerPro-----LysCysValSerGlnThrHisPrometIleGluGlyLysIle	506	
322	ATACTTTCCCCAAAGAAGAGT-----GGAGCT-----	351	
507	LeuTyrAspArgLysLysArgSerLeuProGlyArgAlaValArgGluTyrValAspAsp	526	
352	-----CACAGATTTGATCCATCTCTGTTGTGAAGTAAT-----	384	
527	GluLeuSerThrHisArgGlnHisSerGlyLysCysGlyIleValSerGlyLysLeuGlu	546	
385	-----TGTCACCATGGAACTTCAGTT	405	
547	ArgMetIleMetGlnHisSerAspAenGlyValSerVal	559	

RESULT 12
E0614
probable competence-related protein STY0984 [imported] - Salmonella enterica subsp. enta
Species: Salmonella enterica subsp. enterica serovar Typhi
Note: This species has also been called Salmonella typhi
Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
Accession: AE0614
Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
h, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.

C;Accession: E81457; T02809
R;Myler, P.J.; Audleman, L.; deYos, T.; Hixson, G.; Kiser, P.; Lemley, C.; Magnes, C.;
Proc. Natl. Acad. Sci. U.S.A. 96, 2902-2906, 1999
A;Title: Leishmania major Friedlin chromosome 1 has an unusual distribution of protein-
A;Reference number: A81455; MUID:99178987; PMID:10077609
A;Accession: E81457
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-335 <P>
A;Cross-references: GB:AE001274; NID:g3264850; PIDN:AAC24632.1; PID:g2995585; GSPDB:GNO
A;Experimental source: strain MHOM/IL/81/Friedlin
C;Genetics:
A;Gene: PCTIB
A;Map position: 1
C;Superfamily: hypothetical protein PCTIB; cyclophilin homology
C;Keywords: cis-trans-isomerase
F;149-331/Domain: cyclophilin homology <CYP>

Alignment Scores:
Pred. No.: 2.25 Length: 335
Score: 82.00 Matches: 33
Percent Similarity: 35.94% Conservative: 13
Best Local Similarity: 25.78% Mismatches: 28
Query Match: 11.01% Indels: 54
DB: 1 Gaps: 8

US-09-518-842-1 (1-420) x T02809 (1-335)

QY 251 TCTGATGCTGACCTAAGCGTTGTCATCTTTG-----TTGTTGGAGTTGACACCAT 198
Db 172 AlaSerValAlaProLeuThrCysIysAsnPheCysGluLeuCysArgGlyThrVal 191
QY 197 TCTTTGGAGCTCCAGATCCAGATCCAGCCAC-----CCTCCT----- 162
Db 192 AlaMetGly-----AspThrValAlaHisSerSerValPheAspProProAlaGln 209
QY 162 ----- 162
Db 210 HisIleGlyTyrLysGlyThrThrPhePheArgThrLeuLysAspAlaTrpValMetGly 229
QY 161 GGGTGGTGGTGAATGCTTCTTCAGGCAATGAGGCAATATGAC-----AGCAAGTGT 108
Db 230 GlyAspValThrGlyAlaHisSerGlyAsnGlyTyrSerCysTyrArgCysPhe 249
QY 107 CCA-----AATCGG 99
Db 250 ProAspGluThrTyrAlaValProHisAspAlaGlyIleLeuGlyMetCysAsnAsp 269
QY 98 GGACCACTCCCTCAGCTCTGCTAGGCTTTCTCTAAGG-----AGTTGGCTCAGC 45
Db 270 GlyProHisThrSerSerAlaPheTyrIleThrLeuArgProMetSerTrpMetAsn 289
QY 44 AGCAGCCAGATTGCT---GGCAGA 24
Db 290 GlyLysTyrValAlaPheGlyArg 297

Search completed: October 9, 2003, 12:19:46
Job time : 36 secs

A;Residues: 1-321 <STO>
A;Cross-references: GB:AE004733; GB:AE004091; NID:g9949194; PIDN:AAG06477.1; GSPDB:GNO01
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA3089
C;Superfamily: Neisseria meningitidis hypothetical protein NMB1349

Alignment Scores:
Pred. No.: 1.75 Length: 321
Score: 83.00 Matches: 17
Percent Similarity: 61.54% Conservative: 7
Best Local Similarity: 43.59% Mismatches: 6
Query Match: 10.72% Indels: 9
DB: 2 Gaps: 2

JS-09-518-842-1 (1-420) x E81259 (1-321)

QY 94 GGTCCCGATTGGAAACACTGTGTCAT-----ATTGCCCATG 135
Db 234 GlyProGlnThrGlyGlnHisTrpLeu-HisLeuSerArgGluHisTrpLeuAlaProAl 253
QY 136 CTTGAGAAACATTCCACCA-----CCACCCAGAGGGTGGCTGCTGGAATCTG 184
Db 253 aMetArgArgGlnAspProLeuAspProProArgAspGlyAlaCysTrpAsnIle 271

RESULT 14
T09271
probable tail component protein 1626 - Streptococcus thermophilus phage Sfil9
Species: Streptococcus thermophilus phage Sfil9
Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 11-May-2000
Accession: T09271
Desiere, F.; Lucchini, S.; Brussow, H.
Molecular Biology 241, 345-356, 1998
Title: Evolution of Streptococcus thermophilus bacteriophage genomes by modular exchan
Reference number: 216607; MUID:98160788; PMID:9499809
Accession: T09271
Status: translated from GB/EMBL/DBJ
Molecule type: DNA
Residues: 1-1626 <DES>
Cross-references: ENBL:AF032122; NID:g2935682; PID:g2935689
Keywords: tail protein

Alignment Scores:
red. No.: 2.36 Length: 1626
core: 82.50 Matches: 23
Percent Similarity: 36.25% Conservative: 6
Best Local Similarity: 28.75% Mismatches: 20
Query Match: 10.66% Indels: 31
B: 2 Gaps: 3

S-09-518-842-1 (1-420) x T09271 (1-1626)

Y 9 CTTGTTCCG-----GTCTATCTGCC 29
b 1074 ProValProGlyArgAsnThrTyrAlaValLeuProAlaGlyAlaGluValLeuAsnAla 1093
Y 30 AGCAATCTGGCTGCTGCTGAGCCAACTCTTAGAGAAGAGCCCTAGCAGCAGCTGAGGGG 89
b 1094 SerGluTrpAlaAlaLeuSerGlyAla-----LysProPheAlaLysGlyThrGly 1110
Y 90 ATGTGGTCCCGATTGGAAACACTGCTGCTGGAATCTGGAATCTGAGTCCCAAGAAATGTGTCAAC 149
b 1111 PheTrpSerLysIleTrpAsnThrThr----- 1120
Y 150 CACCACACCCAGAGGGTGGCTGCTGGAATCTGGAATCTGAGTCCCAAGAAATGTGTCAAC 209
b 1121 -----AsnValAlaGlySerValTrpAsnGlyLeuLysAsnGlyValAsp 1135

RESULT 15
02809
robable peptidylprolyl isomerase (EC 5.2.1.8) PCTIB [similarity] - Leishmania major (S
Species: Leishmania major
Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 01-Sep-2000

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

M nucleic - protein search, using frame_plus_n2p model

un on: October 9, 2003, 12:18:46 ; Search time 51 Seconds
(without alignments)
2653.880 Million cell updates/sec

title: US-09-518-842-1

effect score: 774

sequence: 1 ATGCCAGCCTGTTCCGGC.....CACTTAATTATGTACATAG 420

coring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

searched: 600653 seqs, 161128416 residues

total number of hits satisfying chosen parameters: 1201306

inimum DB seq length: 0

aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

command line parameters:

MODEL=frame+ n2p model -DEV=xlp

Q=/cgn2_1/USFO_spool/US0951842/runat_09102003_111040_13189/app query.fasta_1.583

DB=Published Applications_AA -OFMT=fastan -SUFFIX=rapb -MINMATCH=0.1

LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62

TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100

THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0

MAXLEN=2000000000 -USER=US0951842 @CIGN 1.1.83 @runat_09102003_111040_13189

NCPU=6 -ICPU=3 -NO_WMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100

LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=10 -XGAPOP=10 -XGAPEXT=0.5

FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

atabase :

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*

2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*

3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*

4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*

5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*

6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*

7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*

8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*

9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*

10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*

11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*

12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*

13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*

14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*

15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*

16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*

17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*

18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

result	Query	Match	Length	DB	ID	Description
--------	-------	-------	--------	----	----	-------------

	1	228	29, 5	41	12	US-10-339-740-223	Sequence 223, App
	2	147	19.0	25	12	US-10-339-740-224	Sequence 224, App
	3	119.5	15.4	185	15	US-10-205-823-349	Sequence 349, App
	4	78.5	10.1	890	14	US-10-028-056-3	Sequence 3, Appli
	5	76	10.2	619	10	US-09-976-059-30	Sequence 30, Appl
	6	74.5	10.0	1166	15	US-10-101-464A-900	Sequence 900, App
	7	74.5	9.6	1531	12	US-09-841-260-98	Sequence 98, Appl
	8	74.5	9.6	1531	14	US-10-007-693-98	Sequence 98, Appl
	9	74	9.9	857	12	US-10-032-585-7857	Sequence 7857, Ap
	10	74	9.6	1852	12	US-10-023-649-2	Sequence 2, Appli
	11	73.5	9.5	269	9	US-09-815-242-5429	Sequence 5429, Ap
	12	73.5	9.5	273	9	US-09-815-242-12487	Sequence 12487, A
	13	73.5	9.5	273	9	US-09-815-242-12747	Sequence 12747, A
	14	73.5	9.5	274	10	US-09-925-637-12	Sequence 12, Appl
	15	73.5	9.5	274	15	US-10-084-205-12	Sequence 12, Appl
	16	73.5	9.5	282	12	US-09-772-598-1	Sequence 1, Appli
	17	73	9.8	428	15	US-10-001-254-22	Sequence 22, Appl
	18	73	9.4	458	9	US-09-826-463-1	GENERAL INFORMA
	19	73	9.8	560	10	US-09-978-295A-420	Sequence 420, App
	20	73	9.8	560	10	US-09-978-697-420	Sequence 420, App
	21	73	9.8	560	10	US-09-978-192A-420	Sequence 420, App
	22	73	9.8	560	10	US-09-999-832A-420	Sequence 420, App
	23	73	9.8	560	10	US-09-978-189-420	Sequence 420, App
	24	73	9.8	560	11	US-09-978-608A-420	Sequence 420, App
	25	73	9.8	560	11	US-09-978-585A-420	Sequence 420, App
	26	73	9.8	560	11	US-09-978-191A-420	Sequence 420, App
	27	73	9.8	560	11	US-09-978-403A-420	Sequence 420, App
	28	73	9.8	560	11	US-09-978-564A-420	Sequence 420, App
	29	73	9.8	560	11	US-09-999-833A-420	Sequence 420, App
	30	73	9.8	560	11	US-09-981-915A-420	Sequence 420, App
	31	73	9.8	560	11	US-09-978-824-420	Sequence 420, App
	32	73	9.8	560	11	US-09-918-585A-420	Sequence 420, App
	33	73	9.8	560	11	US-09-978-423A-420	Sequence 420, App
	34	73	9.8	560	11	US-09-978-193A-420	Sequence 420, App
	35	73	9.8	560	11	US-09-999-830A-420	Sequence 420, App
	36	73	9.8	560	11	US-09-978-757A-420	Sequence 420, App
	37	73	9.8	560	11	US-09-978-187B-420	Sequence 420, App
	38	73	9.8	560	11	US-09-978-643A-420	Sequence 420, App
	39	73	9.8	560	12	US-09-978-375A-420	Sequence 420, App
	40	73	9.8	560	12	US-09-978-188A-420	Sequence 420, App
	41	73	9.8	560	12	US-09-978-298A-420	Sequence 420, App
	42	73	9.8	560	12	US-10-143-031A-420	Sequence 420, App
	43	73	9.8	560	12	US-10-002-967A-420	Sequence 420, App
	44	73	9.8	560	12	US-10-017-083A-420	Sequence 420, App
	45	73	9.8	560	12	US-10-143-030A-420	Sequence 420, App

ALIGNMENTS

RESULT 1

US-10-339-740-223

; Sequence 223, Application US/10339740

; Publication No. US20030187246A1

; GENERAL INFORMATION:

; APPLICANT: Doherty, Stephen

; APPLICANT: Reddy, Bindu

; APPLICANT: Platt, Darren

; APPLICANT: Ferguson, Kimberly

; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF C. ELEGANS INSULIN-LIKE GENES AND

; FILE OF INVENTION: THEREOF

; FILE REFERENCE: 7326-069-999

; CURRENT APPLICATION NUMBER: US/10339,740

; CURRENT FILING DATE: 2003-01-09

; PRIOR APPLICATION NUMBER: US/09/084,303A

; PRIOR FILING DATE: 1998-05-26

; NUMBER OF SEQ ID NOS: 298

; SOFTWARE: Patent in version 3.1

; SEQ ID NO 223

; LENGTH: 41

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-339-740-223


```
Alignment Scores:
Pred. No.: 6.02e-17 Length: 41
Score: 228.00 Matches: 41
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 29.46% Indels: 0
DB: 12 Gaps: 0

US-09-518-842-1 (1-420) x US-10-339-740-223 (1-41)
QY 52 CAACCTCTTAGAAGAGCTAGCAGCAGCTGAGGGATGTGGTCCCGATTGGAAAA 111
Db 1 GlnLeuLeuArgGluSerLeuAlaGluLeuArgGlyCysGlyProArgPheGlyLys 20
QY 112 CACTTGCTGCATATGCCCCCATGCTGAGAGACATTCCACCACCCAGGAGGTGG 171
Db 21 HisLeuLeuSerTyrCysProMetProGluLysThrPheThrThrProGlyGlyTrp 40
QY 172 CTG 174
Db 41 Leu 41

RESULT 2
US-10-339-740-224
; Sequence 224, Application US/10339740
; Publication No. US20030187246A1
; GENERAL INFORMATION:
; APPLICANT: Doberstein, Stephen
; APPLICANT: Reddy, Bindu
; APPLICANT: Platt, Darren
; APPLICANT: Ferguson, Kimberly
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF C. ELEGANS INSULIN-LIKE GENES AND U
; FILE REFERENCE: 7326-069-999
; CURRENT APPLICATION NUMBER: US/10/339,740
; PRIOR FILING DATE: 2003-01-09
; PRIOR APPLICATION NUMBER: US/09/084,303A
; PRIOR FILING DATE: 1998-05-26
; NUMBER OF SEQ ID NOS: 298
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 224
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-339-740-224

Alignment Scores:
Pred. No.: 6.53e-08 Length: 25
Score: 147.00 Matches: 25
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.99% Indels: 0
DB: 12 Gaps: 0

US-09-518-842-1 (1-420) x US-10-339-740-224 (1-25)
QY 343 AGTGGACGTCACAGATTGATCCATTCTGTGTGAAGTAATTGTGACGATCGAACTTCA 402
Db 1 SerGlyArgHisArgPheAspProPheCysCysGluValIleCysAspAspGlyThrSer 20
QY 403 GTTAAATTATGTACA 417
Db 21 ValIysLeuCysThr 25

RESULT 3
US-10-205-823-349
; Sequence 349, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
```

```

; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Wonsey, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; PRIOR FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 349
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-205-823-349

Alignment Scores:
Pred. No.: 0.000137 Length: 185
Score: 119.50 Matches: 52
Percent Similarity: 37.11% Conservative: 20
Best Local Similarity: 26.80% Mismatches: 57
Query Match: 15.44% Indels: 65
DB: 15 Gaps: 7

US-09-518-842-1 (1-420) x US-10-205-823-349 (1-185)
QY 1 ATGGCCAGCGCTTTCGGTCTTATCTGCAGCAATCTGGCTGCTGAGCCAACTCCTT 60
Db 1 MetProArgLeuPheLeuPheHisLeuLeuGluPheCysLeuLeuLeuAsnGlnPheSer 20
QY 61 AGAGAAAGCCTAGCAGCAGAG-----CTGAGGGGATGTGTCCTCCCGATT 105
Db 21 Arg---AlaValAlaAlaLysTrpLysAspAspValIleLysLeuCysGlyArgGluLeu 39
QY 106 GGAAGAACACTTGTCTATATTGCCCATG-----CCTGAG 141
Db 40 ValArgAlaGlnIleAlaIleCysGlyMetSerThrTrpSerLysArgSerLeuSerGln 59
QY 142 AGACATTCACACACACCCAGGAGGTGGCTGGATCTGGACGTCCC-----AAAGAA 198
Db 60 GluAspAlaProGlnThrPro-----ArgProValAlaGlu 71
QY 199 ATGGTGTCAACCTCCAAACAAAGATGACAAAGCCTTAGGTACGACATCAGATTCA 258
Db 72 IleValProSerPheIleAsnLysAspThrGluThrIleIleMetLeuGluPheIle 91
QY 259 CCTAATTTGTCCAGAGCTGAAGAAACCACTGTCTGAAGGGAGCCCATTTG----- 312
Db 92 AlaAsnLeuProGluLeuLysAlaAlaLeuSerGluArgGlnProSerLeuProGlu 111
QY 313 -----AAGAA 318
Db 112 LeuGlnGlnTyrValProAlaLeuLysAspSerAsnLeuSerPheGluGluPheLys 131
QY 319 ATAATATCTTCCCGC----- 333
Db 132 LeuIleArgAsnArgGlnSerGluAlaAlaAspSerAsnProSerGluLeuLysTyrLeu 151
```

334 -----AAAAGAGAGTGGAGCTCACAGATTGATCCATTCTGT 372
152 GlyLeuAspThrHisSerGlnLysLeuArgProTyrValAlaLeuPheGluLysCys 171
373 TGTGAAGTAATTTGTGAGGATGGAACTTCAGTTAAATATGT 414
172 CysLeuIleGlyCysThrLysArgSerLeuAlaLysTyrCys 185

RESULT 4
US-10-028-056-3
Sequence 3, Application US/10028056
Publication No. US20020152483A1
GENERAL INFORMATION:
APPLICANT: PTERFY, MIKLOS
TITLE OF INVENTION: A NOVEL GENE ASSOCIATED WITH REGULATION OF ADIPOSYTY AND INSULIN
CURRENT APPLICATION NUMBER: US/10/028,056
PRIOR FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: US 60/257,772
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3
LENGTH: 890
TYPE: PRT
ORGANISM: Homo sapiens
US-10-028-056-3

Alignment Scores:
red. No.: 8.39 Length: 890
core: 78.50 Matches: 28
Percent Similarity: 40.68% Conservative: 20
Best Local Similarity: 23.73% Mismatches: 59
Query Match: 10.14% Indels: 11
Gaps: 3

S-09-518-842-1 (1-420) x US-10-028-056-3 (1-890)

34 ATCTGGTGTCTGCTGACCACTCTTAGAGAAAGCTAGCAGAGCTGAGGGATGT 93
274 LeuTrpGlyGluLeuProGlnAlaLysSerSerProHisLysMetLysGluSer 293
94 GGTCCTCCGATTTGGAAA-----CACTTGCTGTCATATTGCCCATG 135
294 SerProLeuSerSerArgLysIleCysAspLysSerHisPheGlnAlaIleHisSerGlu 313
136 CCTGAGAGACATTCACC-----ACCACCCAGAGGGTGGCTGCTGGAACT 183
314 SerSerAspThrPheSerAspGlnSerProThrLeuValGlyGlyAlaLeuLeuAspGln 333
184 GGAGCTCCCAAGAAATGCTGTCACTCCCAACAAAGATGCAAGCTTAGGTACG 243
334 AsnLysProGlnThrGluMetGlnPheValAsnGluLysPleuGluThrLeuGlyAla 353
244 ACATCAGATTCATTCCTAATTTGTCACAGAGCTGAAGAAACCACTGCTGAAGGGCAG 303
354 AlaAlaProLeuLeuProMetIleGlu---GluLeuLysProProSerAlaSerValVal 372
304 CCATCAITGAGAAATAATACTTTCGGCAAAAGAGAGTGGACGCTCAGAC 357
373 GlnThrAlaAsnLysThrAspSerProSerArgLysArgAspLysArgSerArg 390

RESULT 5
S-09-976-059-30
Sequence 30, Application US/09976059
Patent No. US20020164747A1
GENERAL INFORMATION:
APPLICANT: Farnet, Chris
APPLICANT: Zazopoulos, Emmanuel
APPLICANT: Staffa, Alfredo
TITLE OF INVENTION: Genes and Proteins for Biosynthesis of Ramoplanin.
FILE REFERENCE: 3019-PCT

CURRENT APPLICATION NUMBER: US/09/976,059
CURRENT FILING DATE: 2001-10-15
NUMBER OF SEQ ID NOS: 34
SOFTWARE: PatentIn version 3.0
SEQ ID NO 30
LENGTH: 619
TYPE: PRT
ORGANISM: Actinoplanes sp.
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(1)
OTHER INFORMATION: V represents a non-standard initiator codon. It is expected tha
OTHER INFORMATION: the biosynthesized protein will have a formylmethionine residue
US-09-976-059-30

Alignment Scores:
Pred. No.: 14.5 Length: 619
Score: 76.00 Matches: 34
Percent Similarity: 39.82% Conservative: 11
Best Local Similarity: 30.03% Mismatches: 32
Query Match: 10.20% Indels: 36
Gaps: 8

US-09-518-842-1 (1-420) x US-09-976-059-30 (1-619)

QY 253 ATTCTGATGTCGTACCTA-----AGGCTTGTCAT--- 224
Db 214 LeuValValThrTyrLeuAlaGlyAlaProValArgAlaArgAlaValHisVal 233
QY 223 -----CTTTGTTGAGGTTGACACCATTTCTTTGGAGCTCCAGATTCCAGCA 173
Db 234 AlaAlaAlaValAlaAlaThrLeuAlaValSerLeuLeuTrp-----ValLeuAla 250
QY 172 GCCACCTCTCGGGTGGTGGTAATGCTTCTCAGGCATGG----- 131
Db 251 LeuThrLeuLeuProGly-----SerHisArgProTrpAlaAspGlyThrTrpSer 267
QY 130 GCAATATGACAGCAAGTCTTTTC-----CAAATGGGGACACATC 89
Db 268 GlyAsnAlaPheAlaMetValPheGlyTyrAsnGlyPheAspArgAlaGlyIleHisVal 287
QY 88 CCTCAGCTCTCTGCTAGGCTTCTCTAAGCA-----GTTGGCTCAGCAGCAGCC 38
Db 288 ProGlyAlaLeuThrGlyPheThrAspGlyGlyAlaAlaAlaGlyGlySerTrpThr 307
QY 37 ACATTCCTGGCAGATAGGACCGAACA-----GGCTGG 5
Db 308 AlaLeuAlaAlaAspArgLeuAlaThrGlnIleGlyTrp 320

RESULT 6
US-10-101-464A-900
Sequence 900, Application US/10101464A
Publication No. US20030046728A1
GENERAL INFORMATION:
APPLICANT: Strabala, Timothy
APPLICANT: Nieuwenhuizen, Nicolaas
APPLICANT: Higgins, Colleen M.
TITLE OF INVENTION: Compositions Isolated from Plant Cells
TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
FILE REFERENCE: 11000.1020c2
CURRENT APPLICATION NUMBER: US/10/101,464A
CURRENT FILING DATE: 2002-03-18
PRIOR APPLICATION NUMBER: 09/704,302
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 09/228,986
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/162,866
PRIOR FILING DATE: 1999-11-01
PRIOR APPLICATION NUMBER: PCT/US00/00724
NUMBER OF SEQ ID NOS: 989
SOFTWARE: FastSeq for Windows Version 4.0

```

; SEQ ID NO 900
; LENGTH: 1166
; TYPE: PRT
; ORGANISM: Pinus radiata
US-10-101-464A-900

Alignment Scores:
Pred. No.: 25.4 Length: 1166
Score: 74.50 Matches: 39
Percent Similarity: 41.30% Conservative: 18
Best Local Similarity: 28.26% Mismatches: 60
Query Match: 10.00% Indels: 21
Gaps: 15

US-09-518-842-1 (1-420) x US-10-101-464A-900 (1-1166)
QY 413 CATAATTTAACTGAAGTTCATCGTCACAAATTTACTTCAACAGAAATGGATCAATCTG 354
Db 164 AaAsnLeuThrGlyThrileProSerGlyLeuGlnLeuGlnAsnLeu 183
QY 353 TGACGTCCACTTCTCTTTTGGCGGAAAGTATATTTTC-----TTCAATGATGGCTGC 300
Db 184 -----SerLeuSerPheAsnGlyLeuThrGlyLysVal 194
QY 299 CATTACAGACAGTGGTTTCTTCAGCTCTGGTGAACAATAGGAATGAATTTCTGATGCTGA 240
Db 195 ProGluThrPheGlyAsnCyseSerAlaLeuGluLeuGlyLeuGlySerAsnPheLeu 214
QY 239 CTTAAGGCTTGTCCATCT-----TTGTTTGGAGGTTGACACCAATTTCTTTTGGGA 189
Db 215 SerGlyThrileProSerHisLeuGlyMetLeuAlaAaArgileLysileLeuSerLeuGly 234
QY 188 CPTCCAGATTCCAGCAGCCACCTCTGGTGGTGGTGAATGTTCTTCAGGCATGGGG 129
Db 235 ArgAsnGlnLeuSerGlyGlyIleProProSerLeuLeuAsnCysThrGluLeuArgGly 254
QY 128 CAATATGACAGCAAGTGTITTCAAATCGG-----GGACCACTCCCTCAGCTCTGCT 75
Db 255 LeuTyr-----LeuTyrGluAsnArgLeuThrGlyGluIleProTrpGluIleGly 271
QY 74 GCTAGGCTTCTCTAAGGAGTGGCTCAGC-----AGCAGCAGATTCCTGCTGGC 27
Db 272 AlaLysLeuSerLysileGluThrLeuSerLeuGlyAsnGlnLeuSerGly 289

RESULT 7
; Sequence 98, Application US/09841260
; Publication No. US2003017500A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND DIAGNOSIS
; TITLE OF INVENTION: OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.515
; CURRENT APPLICATION NUMBER: US/09/841,260
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 140
; SEQ ID NO 98
; LENGTH: 1531
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis serovar D
JS-09-841-260-98
; Sequence 98, Application US/09841260
; Publication No. US2003017500A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND DIAGNOSIS
; TITLE OF INVENTION: OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.515
; CURRENT APPLICATION NUMBER: US/09/841,260
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 140
; SEQ ID NO 98
; LENGTH: 1531
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis serovar D
JS-09-841-260-98

Alignment Scores:
Pred. No.: 27.3 Length: 1531
Score: 74.50 Matches: 29
Percent Similarity: 44.79% Conservative: 14
Best Local Similarity: 30.21% Mismatches: 26
Query Match: 9.63% Indels: 27
Gaps: 12

US-09-518-842-1 (1-420) x US-09-841-260-98 (1-1531)
QY 79 GAGCTGAGGGGATGGTCCCGATTTGGAAACACTTGT----- 117
Db 472 GluTyrGlnGlyGlyAlaLeuPheGlyGluAsnIleSerLeuSerGluAsnAlaGly 491
QY 118 ----CTGTCATATTGCCCATGCTGCTGAGAGACATTCCACCACC-----CCA 162
Db 492 ValLeuThrPheLysAspAsnIleValIlysthrPheAlaSerAsnGlyLysileLeuGly 511
QY 163 GGAGGGTGGCTGCTGGAATCTGCACCTCCCAAGAAATGGTGTCAACCTCCCAACACAAA 222
Db 512 GlyGlyAlaIleLeuAlaThrGlyLys-----ValGluIleThrAsnAsnSer 527
QY 223 GATGGA-----CAAGCCTTAGGTACGACATCAGAA 252
Db 528 GluGlyIleSerPheThrGlyAsnAlaArgAlaProGlnAlaLeuProThrGlnGluGlu 547
QY 253 TTTATTCCTTAATTTGTCCACGAGCTGAAGAACCACTGTCTGAGGG 300
Db 548 Phe---ProLeuPheSerLysLysGluGlyArgProLeuSerSerGly 562

RESULT 8
; Sequence 98, Application US/10007693
; Publication No. US20020146776A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
; TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.515C2
; CURRENT APPLICATION NUMBER: US/10/007,693
; CURRENT FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 157
; SEQ ID NO 98
; LENGTH: 1531
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis serovar D
US-10-007-693-98
; Sequence 98, Application US/10007693
; Publication No. US20020146776A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
; TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.515C2
; CURRENT APPLICATION NUMBER: US/10/007,693
; CURRENT FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 157
; SEQ ID NO 98
; LENGTH: 1531
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis serovar D
US-10-007-693-98

Alignment Scores:
Pred. No.: 27.3 Length: 1531
Score: 74.50 Matches: 29
Percent Similarity: 44.79% Conservative: 14
Best Local Similarity: 30.21% Mismatches: 26
Query Match: 9.63% Indels: 27
Gaps: 5

US-09-518-842-1 (1-420) x US-10-007-693-98 (1-1531)
QY 79 GAGCTGAGGGGATGGTCCCGATTTGGAAACACTTGT----- 117
Db 472 GluTyrGlnGlyGlyAlaLeuPheGlyGluAsnIleSerLeuSerGluAsnAlaGly 491
QY 118 ----CTGTCATATTGCCCATGCTGCTGAGAGACATTCCACCACC-----CCA 162
Db 492 ValLeuThrPheLysAspAsnIleValIlysthrPheAlaSerAsnGlyLysileLeuGly 511
QY 163 GGAGGGTGGCTGCTGGAATCTGCACCTCCCAAGAAATGGTGTCAACCTCCCAACACAAA 222
Db 512 GlyGlyAlaIleLeuAlaThrGlyLys-----ValGluIleThrAsnAsnSer 527
QY 223 GATGGA-----CAAGCCTTAGGTACGACATCAGAA 252
Db 528 GluGlyIleSerPheThrGlyAsnAlaArgAlaProGlnAlaLeuProThrGlnGluGlu 547
QY 253 TTTATTCCTTAATTTGTCCACGAGCTGAAGAACCACTGTCTGAGGG 300
Db 548 Phe---ProLeuPheSerLysLysGluGlyArgProLeuSerSerGly 562

RESULT 9
; Sequence 7857, Application US/10032585
US-10-032-585-7857
; Sequence 7857, Application US/10032585

```

Alignment Scores:

Pred. No.:	32.7	Length:	1852
Score:	74.00	Matches:	32
Percent Similarity:	40.69%	Conservative:	27
Best Local Similarity:	22.07%	Mismatches:	40
Query Match:	9.56%	Indels:	46
DB:	12	Gaps:	6

US-09-518-842-1 (1-420) x US-10-023-649-2 (1-1852)

QY	4	GCACGCTGTTCCGGTCTCTATCTGCAGCAATCTGGCTGCTGCTGAGCCAACTCTCTTAGA	63
Db	501	AlaGluGluIleArgGluAlaLeuProPro-	517
		-----LeuLeuProGluIleGluLys	
QY	64	GAAGACCTTAGCAGCAGAGCTG-	105
Db	518	GluThrValGluAlaGluValAspLeuIleMetGlnGluAlaGlyAlaGlySer-	535
QY	106	GGAAAAACACTTGTGTCATATTGCCCATGCTGAGAGACATTCACCAACCCAGCA	165
		-----ValGluThrProArg	540
Db	536	-----	
QY	166	GGGTGGCTGCGAATCTGGACGTCCTCAAGAAATGCTGCACTCCACACAAAGAT	225
Db	541	GlyHisIleArgValThrSerTyrProGlyGlu-	551
QY	226	GCACAAAGCCTTAGGTAGACACATCAGAAATCAITTCCTAAATTTGTCACGAGCTGAAAGAA	285
Db	552	---GluLysIleGlySerTyrAlaIleLeuSerProGlnAlaValLeuAsnSerGluLys	570
QY	286	-----CCACTGCTGAGAGCGCCATCATTTGAAGAAAATAATACTTCTCC	330
Db	571	LeuAlaCysIleHisProLeuAlaGlu-	585
QY	331	CGCAAAAGAGAACTGGACGTCACAGATTTGATCCATTCGTTGTGAAGTAATTTCTGAC	390
Db	586	HisLysGlyArgAlaGlyArgTyrLysValGluProTyrHisGlyLysValIleValPro	605
QY	391	GATGAACTTCAGTT	405
Db	606	GluGlyThrAlaVal	610

RESULT 11

US-09-815-242-5429

; Sequence 5429 Application US/09815242

; Patent No. US20020061569A1

GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; PRIOR FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

```

, TITLE OF INVENTION: PROKARYOTES
, FILE REFERENCE: ELITRA_011A
, CURRENT APPLICATION NUMBER: US/09/815,242
, CURRENT FILING DATE: 2001-03-21
, PRIOR APPLICATION NUMBER: 60/191,078
, PRIOR FILING DATE: 2000-03-21
, PRIOR APPLICATION NUMBER: 60/206,848
, PRIOR FILING DATE: 2000-05-23
, PRIOR APPLICATION NUMBER: 60/207,727
, PRIOR FILING DATE: 2000-05-26
, PRIOR APPLICATION NUMBER: 60/242,578
, PRIOR FILING DATE: 2000-10-23
, PRIOR APPLICATION NUMBER: 60/253,625
, PRIOR FILING DATE: 2000-11-27

```

ORGANISM: H0
S-09-925-637-12

Qy	76	GCAGAGCTGAGGGGATGTGGTCCCGATTATTGGAAAAACACTTGTCTGCATATTCTGCCCATG	135
Db	182	AlaProIlePheGlyLeuAsnLysArgGlnGlyArgGlnLeuLeuAlaTyrLeuGlyAla	201
Qy	136	CCTGAGAGACATATCCACCACCCGAGGAGGGTGGCTGCTGGAATCTTGGACGTCCCAA	195

```

Db      202 ProLysGluLeuTy-GluLysThrProThrAla---AspLeuGluAspAspLysProGln 220
QY      196 GAATGGTGTCAACCTCCAAACAACAAAGATGACAAAGCCTTAGGTACGACATCAGAATTC 255
Db      221 -----LeuProAspGluAspAlaLeuGlyValThrTyGluAla 233
QY      256 ATTCCTAAT-----TTGTCACGAG-----CTGAGAAA 285
Db      234 IleAspAsnTyLeuGluGlyLysProValThrProGluGluGlnLysValIleGlu-As 253
QY      286 CCACTGTCTGAAGGCGCCATCATTTGAGAAATAATACTTCCCGCAAAAGAGAGT 345
Db      253 rHISTYrIleArgAsnAlaHis-----LysArgGluLe 264
QY      346 GGACGTCACA 355
Db      264 uAlaTyThr 267

```

Search completed: October 9, 2003, 12:32:15
 Job time : 55 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

DM nucleic - protein search, using frame_plus_n2p model

Run on: October 9, 2003, 12:08:36 : Search time 61.5 Seconds
(without alignments)
2167.973 Million cell updates/sec

Title: US-09-518-842-1

Perfect score: 774

Sequence: 1 ATGCCAGCCTGTTCCGTC.....CAGTTAAATTATGACATAG 420

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 2215726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

MODPT=frame+n2p,model -DEV=xlp
Q=/cgn2_1/USPTC/spool/US09518842/runat_09102003_111035_13009/app_query.fasta_1.583
DB=A_Geneseq_19Jun03 -QMT=fabcan -SUFFIX=rag -MINMATCH=0 -LOOPCL=0
LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
USER=US09518842 -CGCN 1 1 137 @runat_09102003_111035_13009 -NCPU=6 -ICPU=3
NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPLOC=100 -LONGLOG
DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=10 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	739	95.5	139	17	Human early placen
2	739	95.5	139	18	Human relaxin-rela
3	739	95.5	139	19	Zinsl protein. Ho
4	739	95.5	139	20	Human insulin-like
5	739	95.5	139	20	Human early placen
6	620	80.1	124	19	Zinsl NF protein.
7	619	80.0	159	19	N-terminally tagge
8	180	23.3	182	4	Sequence of porcine
9	173	22.4	182	5	Sequence of porcine
10	123.5	16.0	185	10	Amino acid sequenc
11	120.5	15.6	185	5	Sequence of human
12	119.5	15.4	185	5	Sequence of human
13	119.5	15.4	185	5	Sequence of human
14	119.5	15.4	185	23	Prostate cancer-as
15	113.5	14.7	150	16	Prolrelaxin fragme
16	113.5	14.7	162	11	H2 prorelaxin gene
17	112.5	14.5	185	10	Amino acid sequenc
18	111.5	14.4	178	14	Squamous cell spec
19	111	14.3	220	22	Angiotensin conver
20	103	13.8	787	22	Novel human diagno
21	103	13.8	787	22	Novel human diagno
22	101.5	13.1	164	11	H2 prorelaxin dedu
23	88.5	11.4	754	22	Human protein sequ
24	88.5	11.4	754	23	Novel human protei
25	80	10.3	239	21	Human insulin-like
26	80	10.3	240	21	A human insulin-li
27	80	10.3	322	13	h-IGFBP-4. Homo s
28	80	10.3	324	13	Sequence of insuli
29	78.5	10.1	912	24	Human secretory po
30	78	10.5	210	22	Proplonibacterium
31	78	10.1	927	22	Novel human diagno
32	76.5	10.3	402	21	Arabidopsis thalia
33	76.5	9.9	468	22	Proplonibacterium
34	76	10.2	212	24	Breast specific re
35	76	10.2	619	23	Human polypeptide,
36	76	9.8	740	22	Human NF-kappaB as
37	76	9.8	812	22	Novel human diagno
38	75.5	9.8	1034	22	Novel human diagno
39	75.5	9.8	1034	22	HMG-CoA reductase
40	75	10.1	994	18	Osteotesticular pr
41	75	10.1	1711	19	Mutant osteotestic
42	75	10.1	1711	19	Human ORF3048 prot
43	74.5	10.0	135	23	Pinus radiata cell
44	74.5	10.0	1166	21	C. trachomatis CT8
45	74.5	9.6	1531	22	

ALIGNMENTS

RESULT 1
AAR89134
ID AAR89134 standard; Protein; 139 AA.
XX
AC AAR89134;
XX
DT 31-JUL-1996 (first entry)
XX
DE Human early placental insulin-like protein.
XX
KW Insulin growth hormone family; early placental insulin like protein;
KW placental; placental tissue; cytotrophoblast; trimester; probe; primer;
KW amplification; polymerase chain reaction; tyrosine phosphorylation;
KW cellular protein; growth factor; human; lactation; promoter; PCR;
KW regeneration; nerve; muscle; skin; bone tissue.
XX
OS Homo sapiens.
XX

FH Key Location/Qualifiers
 FT Peptide 1..17
 FT Peptide /note= "signal peptide"
 FT Peptide 18..58
 FT Peptide /note= "B-chain peptide"
 FT Peptide 59..109
 FT Peptide /note= "C-peptide, links B and A chains in pre-B"
 FT Peptide 110..139
 FT Peptide /note= "A-chain peptide"
 XX W09534653-A1.
 XX 21-DEC-1995.
 XX 12-JUN-1995; 95WO-FR00766.
 XX 13-JUN-1994; 94FR-0007191.
 XX (INSR) INST ROUSSY GUSTAVE.
 XX Bellet D, Chassin D, Koman A;
 XX WPI: 1996-049682/05.
 XX N-PSDB; AAT10275.
 XX Early placental insulin-like protein, EPIL/placentalin - contains
 XX growth factor-like activity useful for e.g. promoting lactation or
 XX for regeneration of nerve, muscle, skin or bone tissue
 XX Claim 1: Page 14-15; 25pp; French.
 XX This is the amino acid sequence of a novel member of the insulin growth
 XX hormone family designated early placental insulin like (EPIL) protein or
 XX placentalin. The encoding gene has been found to be expressed exclusively
 XX in the early placental tissue with a small amount in full term placental
 XX tissue. No other tissue expresses this gene. The gene was obtained from
 XX a cDNA library derived from mRNA isolated from cytotrophoblasts derived
 XX from first trimester placental tissue using, as a probe, a fragment of
 XX the gene amplified by primers AAT10276-7. Although the specific
 XX activity of the protein remains to be elucidated, it is thought that the
 XX protein will induce tyrosine phosphorylation of cellular proteins and may
 XX have growth factor-like activities e.g. human growth factor type 1 or
 XX lactation promoter activities. It may also be used to regenerate e.g.
 XX nerve, muscle, skin or bone tissue.
 XX Sequence 139 AA;
 Alignment Scores:
 Aligned No.: 139 Length: 139
 core: 739.00 Matches: 139
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 95.48% Indels: 0
 B: 17 Gaps: 0
 US-09-518-842-1 (1-420) x AAR89134 (1-139)
 Y 1 ATGCCAGCCTGTTCCGGTCTATCTGCCAGCAATCTGGCTGCTGTGAGCCACCTCCT 60
 b 1 MetAlaSerLeuPheArgSerTyrLeuProAlaIleTrpLeuLeuLeuSerGlnLeuLeu 20
 Y 61 AGAGAAAGCCTAGCAGCAGAGCTGAGGGGATGGTCCCGATTTCGAAACACTTGTG 120
 b 21 ArgGluSerLeuAlaAlaGluLeuArgGlyCysGlyProArgPheGlyLysHisLeuLeu 40
 Y 121 TCATATTCCCATCGCTGAGAGACATTTCACCACCAACCCAGGAGGGTGGCTGTGGAA 180
 b 41 SerTyrCysProMetProGluLysThrPheThrThrThrProGlyGlyTrpLeuLeuGlu 60
 Y 181 TCTGACCTCCCAAGAAATGCTGTCAACCTCCAGACAAAGATGACAGACCTTAGCT 240
 b 61 SerGlyArgProLysGluMetValSerThrSerAsnAsnLysAspGlyGlnAlaLeuGly 80

QY 241 ACGACATCAGAAATTCCTAATTTGTCCACGAGCTGAAGAACCACTGTCTGAAGG 300
 Db 81 ThrThrSerGluPheIleProAsnLeuSerProGluLeuLysLysProLeuSerGluGly 100
 QY 301 CAGCCATCATTGAGAAATAATTAATCTTCGCGCAAAAAGAGAGAGTGCACGATTT 360
 Db 101 GlnProSerLeuLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 120
 QY 361 GATCCATTCTGTTGTGAAGTAATTTGTGACGATCGAACTTCAGTTAAATATGTACA 417
 Db 121 AspProPheCysCysGluValIleCysAspAspGlyThrSerValLysLeuCysThr 139
 RESULT 2
 AAW17676
 ID AAW17676 standard; Protein; 139 AA.
 XX AAW17676;
 XX 24-JUL-1997 (first entry)
 XX Human relaxin-related factor-2 (RRF-2).
 DB Relaxin-related factor-2; RRF-2; testis; sperm; growth factor;
 KW infertility.
 KW Homo sapiens.
 OS Key Location/Qualifiers
 FH Peptide 1..26 /label= Sig_peptide
 FT Misc-difference 124 /note= "conserved Cys residue indicative of
 FT insulin family member"
 FT Misc-difference 125 /note= "conserved Cys residue indicative of
 FT insulin family member"
 FT Misc-difference 129 /note= "conserved Cys residue indicative of
 FT insulin family member"
 FT Misc-difference 138 /note= "conserved Cys residue indicative of
 FT insulin family member"
 FT W09716549-A2.
 XX 09-MAY-1997.
 XX 01-NOV-1996; 96WO-US17342.
 XX 21-FEB-1996; 96US-0012016.
 XX 03-NOV-1995; 95US-0006221.
 XX (REGE-) REGENERON PHARM INC.
 XX Davis S;
 XX WPI: 1997-272118/24.
 XX N-PSDB; AAT68419.
 XX New isolated relaxin-related factor genes - used to develop products
 XX which can be used in diagnosis and therapy, e.g. in fertility and
 XX pregnancy applications
 XX Example 3; Fig 4; 34pp; English.
 XX Human relaxin-related factor-2 (RRF-2) (AAW17676) is a placenta-
 XX specific growth factor related to relaxin and to the insulin family
 XX of ligands. RRF-2 cDNA (AAT68419) was isolated in a search of
 XX expressed sequence tags for sequences related to relaxin. RRF-2
 XX displays all the expected features of a new insulin family member,
 XX partic. with regard to a cluster of four cysteine residues at the
 XX C-terminus of the molecule. A related testis-specific factor,
 XX RRF-1 (AAW17675), has also been identified. RRF-2 can be produced

CC in transformed host cells for use in the prepn. of antibodies and
 CC therapeutic compns., or as a growth factor for maintaining cells
 CC in culture. RRE-2 may be useful for modulating the reproductive
 CC physiology of mammals during pregnancy and parturition.

XX Sequence 139 AA;

Alignment Scores:
 Pred. No.: 1,19e-80 Length: 139
 Score: 739.00 Matches: 139
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 95.48% Indels: 0
 DB: 18 Gaps: 0

US-09-518-842-1 (1-420) x AAU17676 (1-139)

1 ATGGCCAGCCCTGTCGGTCTCTATCTGCAGCAATCTGGCTGCTGCTGAGCCAACTCCTT 60
 1 MetAlaSerLeuPheArgSerTyLeuProAlaIleTrpLeuLeuSerGlnLeuLeu 20
 61 AGAGAAAGCCTAGCAGCAGAGCTGAGGGGATGTGTCCCGATTGGAAAACACTTGCTG 120
 21 ArgGluSerLeuAlaAlaGluLeuArgGlyCysGlyProArgPheGlyLysHisLeuLeu 40
 121 TCATATTGCCCATGCTGCTGAGAGACATTCACCAACCCAGAGGGTGGCTGCTGGAA 180
 41 SerTyCysProMetProGluLeuPheThrThrThrProGlyGlyTrpLeuLeuGlu 60
 181 TCTGGAGCTCCCAAGAAATGGTCTCACTCCCAACAAAGATGGCAAGCTTAGGT 240
 61 SerGlyArgProLysGluMetValSerThrSerAsnAsnLysAspGlyGlnAlaLeuGly 80
 241 ACGACATCAGAAATTCATTCCTAAATTTGTCACAGAGCTGAAGAAACCACTGTCTGAAGGG 300
 81 ThrThrSerGluPheIleProAsnLeuSerProGluLeuLysPheProLeuSerGluGly 100
 301 CAGCCATCATTGAAGAAATAATCTTCCCGCAAAAGAGAGTGGACGCTCAGATTT 360
 101 GlnProSerLeuLysLysIleLeuSerArgLysLysArgSerGlyArgHisArgPhe 120
 361 GATCCATTCTGTTGAAGTAATTTGTGACGATCGAACTTCAGTTAATTTATGTACA 417
 121 AspProPheCysCysGluValIleCysAspAspGlyThrSerValLysLeuCysThr 139

RESULT 3

AW69168
 AAU69168 standard; Protein; 139 AA.

XX

XX

XX 07-OCT-1998 (first entry)

XX Zinsl protein.

XX Zinsl; human; placenta; placentin; pancreatic islet cell proliferation;

XX insulin secretion; diabetes; therapy.

XX Homo sapiens.

XX

XX

XX Key

XX Region

XX /note= "B chain of Zinsl"

XX 115..0

XX /note= "A chain of Zinsl"

XX

XX W09827210-A1.

XX

XX 25-JUN-1998.

XX

XX 16-DEC-1997; 97WO-US23326.

XX

XX 16-DEC-1996; 96US-0033003.

XX

(ZYMO) ZYMOGENETICS INC.

Conklin DC, Hoffmann RC, Humes JM, Jaspers SR, Ren HP;
 PI Sprugel KH;

WPI: 1998-362779/31.
 DR N-PSDB; AAU44663.

New isolated protein, Zinsl - obtained from human placenta, which
 PT increases the proliferation of pancreatic islet cells, used for
 PT treating diabetes

PS Claim 4; Page 60-61; 77pp; English.

CC This sequence is the human Zinsl protein of the invention. The Zinsl
 CC protein was isolated from human placenta, and is believed to be a new
 CC version of the mature protein of placentin, having disulphide bonded A
 CC and B chains. The protein can be used for stimulating the proliferation
 CC of pancreatic islets to increase insulin secretory capacity of mammals.
 CC In particular it can be used for the treatment of diabetes. It can also
 CC be used for stimulating in vitro proliferation of pancreatic islet cells.
 CC It can also be used for production of antibodies and in detection and
 CC diagnosis.

SQ Sequence 139 AA;

Alignment Scores:

Pred. No.: 1,19e-80 Length: 139
 Score: 739.00 Matches: 139
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 95.48% Indels: 0
 DB: 19 Gaps: 0

US-09-518-842-1 (1-420) x AAU69168 (1-139)

1 ATGGCCAGCCCTGTCGGTCTCTATCTGCAGCAATCTGGCTGCTGCTGAGCCAACTCCTT 60
 1 MetAlaSerLeuPheArgSerTyLeuProAlaIleTrpLeuLeuSerGlnLeuLeu 20
 61 AGAGAAAGCCTAGCAGCAGAGCTGAGGGGATGTGTCCCGATTGGAAAACACTTGCTG 120
 21 ArgGluSerLeuAlaAlaGluLeuArgGlyCysGlyProArgPheGlyLysHisLeuLeu 40
 121 TCATATTGCCCATGCTGCTGAGAGACATTCACCAACCCAGAGGGTGGCTGCTGGAA 180
 41 SerTyCysProMetProGluLeuPheThrThrThrProGlyGlyTrpLeuLeuGlu 60
 181 TCTGGAGCTCCCAAGAAATAATCTTCCCGCAAAAGAGATGGACGCTCAGATTTAGGT 240
 61 SerGlyArgProLysGluMetValSerThrSerAsnAsnLysAspGlyGlnAlaLeuGly 80
 241 ACGACATCAGAAATTCATTCCTAAATTTGTGACGATCGAACTTCAGTTAATTTATGTACA 417
 81 ThrThrSerGluPheIleProAsnLeuSerProGluLeuLysPheProLeuSerGluGly 100
 301 CAGCCATCATTGAAGAAATAATCTTCCCGCAAAAGAGAGTGGACGCTCAGATTT 360
 101 GlnProSerLeuLysLysIleLeuSerArgLysLysArgSerGlyArgHisArgPhe 120
 361 GATCCATTCTGTTGAAGTAATTTGTGACGATCGAACTTCAGTTAATTTATGTACA 417
 121 AspProPheCysCysGluValIleCysAspAspGlyThrSerValLysLeuCysThr 139

RESULT 4

AAU26926

ID AAU26926 standard; Protein; 139 AA.

XX

XX

XX AAU26926;

XX

XX 21-DEC-1999 (first entry)

XX

Human insulin-like 4 protein.

Human; insulin-like 4; INS4; embryonic; c bone tissue; ligament;
early-placental insulin-like protein; EPIL; antibody; probe; primer;
diagnosis; pathology; differentiation; proliferation; cartilage;
ossification; osteoporosis; dysplasia.

Homo sapiens.

WO9937780-A1.

29-JUL-1999.

22-JAN-1999; 99WO-FR00137.

23-JAN-1998; 98FR-0000715.

(INSR) INST ROUSSY GUSTAVE.

Laurent A, Bellet D;

WPI; 1999-469135/39.

N-PSDB; AAX83561.

Expression of the INS4 gene in human embryonic bone tissue and

ligaments, -

Disclosure; Fig 1; 60pp; French.

This sequence represents the human protein encoded by the insulin-like 4 (INS4) gene which is designated early-placental insulin-like (EPIL) protein. The INS4 gene is expressed in human embryonic bone tissue and ligaments and encodes 3 different EPIL proteins designated EPIL 1, 2 or 3. EPIL 1 is a single chain comprising amino acids 18-139, EPIL 2 is a 2 chain protein with chain A comprising amino acids 115-139 and chain B comprising amino acids 18-58 and EPIL 3 is a single chain comprising amino acids 59-114. The nucleic acids and protein or antibodies against EPIL 1, 2 or 3, and probes or primers for INS4 are useful for the diagnosis of pathology associated with abnormal differentiation and/or proliferation of bone tissue or ligaments or development of abnormal cartilage and/or abnormal ossification of forming bones, e.g. osteoporosis or dysplasia. Compositions capable of modulation differentiation, regeneration and/or proliferation of bone tissue and/or ligament cells are also useful in treating bone disease.

Sequence 139 AA;

Alignment Scores:
red. No.: 1.19e-80 Length: 139
core: 739.00 Matches: 139
Percent Similarity: 100.00% Conservative: 0
est Local Similarity: 100.00% Mismatches: 0
Query Match: 95.48% Indels: 0
B: 20 Gaps: 0

US-09-518-842-1 (1-420) x AAY26926 (1-139)

Y 1 ATGCCAGCTGTTCGGCTCTATCTGCCAGCAATCTGCTGCTGAGCAACTCTCTT 60
b 1 MetAlaSerLeuPheArgSerTyrLeuProAlaIleTrpLeuLeuLeuSerGlnLeuLeu 20
Y 61 AGAAGAAGCTAGCAGCAGCTGAGGGGATGTGTCCCGATTGGAAACACTTGTCTG 120
b 21 ArgGluSerLeuAlaAlaGluLeuArgGlyCysGlyProArgPheGlyHisLeuLeu 40
Y 121 TCATATTGCCCATGCTGAGAGACATTCCACACCCAGAGGGTGGCTGTGCGAA 180
b 41 SerTyrCysProMetProGluLysThrPheThrThrThrProGlyGlyTrpLeuLeuGlu 60
Y 181 TCTGGACGTCCCAAGAAATGGTGTCAACTCCACACAAAGATGGACAGCCTTAGGT 240
b 61 SerGlyArgProLysGluMetValSerThrSerAsnAsnLysAspGlyGlnAlaLeuGly 80

QY 241 ACGACATCAGATTCATTCTCTAATTGTCCACGAGCTGAAGAAACCACTGTCTGAAGGG 300
Db 81 ThrThrSerGluPheIleProAsnLeuSerProGluLeuLysLysProLeuSerGluGly 100
QY 301 CAGCCATCATTTGAGAAAATAATCTTCCGCCAAAAGAGAGAGTGGACGTACAGATTT 360
Db 101 GlnProSerLeuLysLysIleLeuSerArgLysLysArgSerGlyArgHisArgPhe 120
QY 361 GATCCATTCTGTGTGAAGTAATTGTGACGATGCACTTCAGTTAAATATGTACA 417
Db 121 AspProPheCysCysGluValIleCysAspAspGlyThrSerValLysLeuCysThr 139

RESULT 5
AAW99574

ID AAW99574 standard; Protein; 139 AA.

XX AAW99574;

DT 22-JUN-1999 (first entry)

XX Human early placental insulin-like (EPIL) polypeptide.

DE INS4-4; insulin-like gene; EPIL; early placental insulin-like; antibody;
vector; probe; hybridisation; tumour; hypoglycaemia; hyperglycaemia;
diabetes; cardiovascular.

XX Homo sapiens.

XX WO9909172-A1.

XX 25-FEB-1999.

XX 12-AUG-1998; 98WO-FR01799.

XX 03-NOV-1997; 97FR-0013802.

XX 14-AUG-1997; 97FR-0010387.

XX (INSR) INST ROUSSY GUSTAVE.

XX Bellet D, Bidart JM, Troalen P, Mock P;

XX WPI; 1999-181038/15.

XX N-PSDB; AAX27490.

XX EPIL polypeptides encoded by insulin-like gene 4 - and corresponding
nucleic acids, antibodies, probes, primers, etc.

PS Claim 1; Fig 1; 119pp; French.

XX This sequence represent the early placental insulin-like (EPIL) protein
encoded by the INS4-4 (insulin-like gene 4) gene. The polypeptide,
antibodies to the polypeptide, vectors containing the coding sequence and
probes derived from the coding sequence, can be used to treat tumours,
preferably angioproliferative tumours, especially Kaposi's sarcoma,
tumours of the pancreas, liver, uterus or breast, angiosarcomas,
glioblastomas, neuroblastomas, rhabdomyosarcomas or leiomyosarcomas;
to promote vascularisation of specific tissues; to treat retinopathy,
macular degeneration, psoriasis, endometriosis, rheumatoid arthritis,
atherosclerosis or hyperthyroidism; to treat post-angioplastic
stenosis; to promote or inhibit embryo implantation; to prevent and/or
treat disorders directly or indirectly connected with insulin-like
activity; to prevent and/or treat disorders directly or indirectly
connected with a dysfunction in carbohydrate metabolism, especially
connected with hypo glycaemia or hyperglycaemia, especially gestational
diabetes and diabetic complications, especially cardiovascular
complications.

SQ Sequence 139 AA;

Alignment Scores:
Pred. No.: 1.19e-80 Length: 139
Score: 739.00 Matches: 139
Percent Similarity: 100.00% Conservative: 0

```

Best Local Similarity: 100.00%      Mismatches: 0
Query Match:          95.48%        Indels: 0
DB:                   20            Gaps: 0

JS-09-518-842-1 (1-420) x JAW99574 (1-139)

```

1	ATGGCGACGCTGTTCCGGTCTCTATCTGCGACGCAATCTGGCTGCTGCTGAGCCAACTCCCTT	60
1	MetAlaSerLeuPheArgSerTyrLeuProAlaIleTyrLeuLeuLeuSerGlnLeuLeu	20
61	AGAGAAAGCCTTAGCAGCAGACGCTCAGGGGATGTGGTCCCCGATTTTGGAAAAACACTTGGCTG	120
21	ArgGluSerLeuAlaAlaGluLeuArgGlyCysGlyProArgPheGlyLysHisLeuLeuLeu	40
121	TCATATTATGCCCATTCGCTGAGAACACATTCACCAACCACCCAGAGAGGTGGCTGCTGGAA	180
41	SerTyrCysProMetProGluLysThrPheThrThrThrProGlyGlyTyrLeuLeuLeuGlu	60
181	TCTGGCGTCCCAAGAANAATGGTGTCAACCTCCCAACCAACAAGATGGACACAGGCTTAGCT	240
61	SerGlyArgProLysGluMetValSerThrSerAsnAsnLysAspGlyGlnAlaLeuGly	80
241	ACGACATCAGAAATTCATTTCCTAAATTTGTCACACAGAGCTGAAGAAACCACTGTCTGAAGG	300
81	ThrThrSerGluPheIleProAsnLeuSerProGluLeuLysLysProLeuSerGluGly	100
301	CAGCCATCATTTCAAGAAAAATAATACTTTCCCGCAAAAGAGAAGTGGACGCTCACAGATT	360
101	GlnProSerLeuLysLysIleLeuSerArgLysLysArgSerGlyArgHisArgPhe	420
361	GATCATCTCTGTTGGAAGTAATTTTGTCAGATCGAACTTCAGTTAAATTTATGTACA	417
121	AspProPheCysCysGluValIleCysAspAspGlyThrSerValLysLeuCysThr	139

RESULT 6	
AW69169	
D D	AAW69169 standard; Protein; 124 AA.
X X	
X X	*AAW69169;
X X	
T T	07-OCT-1998 (first entry)
X X	
E X	Zinsl NF protein.
X X	
X W	Zinsl; human; placenta; pancreatic islet cell proliferation;
W W	insulin secretion; diabetes; therapy; Zinsl NF.

This sequence is the human Zins1 NF protein of the invention. The Zins1 protein was isolated from human placenta, and is believed to be a new version of the mature protein of placentin, having disulphide bonded A

CC and B chains. The protein can be used for stimulating the proliferation
CC of pancreatic islets to increase insulin secretory capacity of mammals.
CC In particular it can be used for the treatment of diabetes. It can also
CC be used for stimulating in vitro proliferation of pancreatic islet cells.
CC It can also be used for production of antibodies and in detection and
CC diagnosis.

Sequence 124 AA;

Alignment Scores:		
Pred. No.:	3,148-66	Length: 124
Score:	620.00	Matches: 114
Percent Similarity:	100.0%	Conservative: 1
Best Local Similarity:	99.13%	Mismatches: 0
Query Match:	80.10%	Indels: 0
DB:	19	Gaps: 0

US-09-518-842-1 (1-420) x AAW69169 (1-124)

Qy	73	GCACGAGAGCTGAGGGAGATGTGGTCCCGATTTTGGAAAAACACTTGTCTGCATATATTCGCCCC	132
Db	10	SerAlaGluLeuArGGlYcYsGlyProArgPheGlyLysHisLeuLeuSerTyrCysPro	29
Qy	133	ATGCGCTGAGAAGACATTACACACCACCCAGAGGGGTGGCTGTCTGGAATCTGGACGCTGCC	192
Db	30	MetProGluLysThrPheThrThrThrProGlyGlyTyrLeuLeuGluSerGlyArgPro	49
Qy	193	AAAGAAATGGTGTCAACTCCACACAAAGATGGACAGGCCTTAGGTACGACATCAGAA	252
Db	50	LysGluMetValSerThrSerAsnAsnLysAspGlyGlnAlaLeuGlyThrSerGlu	69
Qy	253	TTCAITTCCTAAATTTGTACACAGAGCTGAAGAAACCACTGTCTGAAGGAGCGCCATCATTTG	312
Db	70	PheIleProAsnLeuSerProGluLeuLysLysProLeuSerGluGlyGlnProSerLeu	89
Qy	313	AGAAAAATAATCTTCCCGAAAAAGAGAGTGGAGCTCAGACATTTGATCCATTCGT	372
Db	90	LysLysIleIleLeuSerArgLysLysArgSerGlyArgHisargPheAspProPheCys	109
Qy	373	TCTGAAGTAATTTGTGACGATGGAACTTCAGTTAAATTATGTACA	417
Db	110	CysGluValIleCysAspAspGlyThrSerValLysLeuCysThr	124

RESULT 7

RESULTS	
AAW69170	
ID	AAW69170 standard; Protein; 159 AA.
XX	
AC	
XX	AAW69170;
XX	
DT	07-OCT-1998 (first entry)
XX	
DE	N-terminally tagged Zinsl protein.
XX	
XX	Zinsl; human; placenta; placentin; pancreatic islet cell proliferation;
KW	insulin secretion; diabetes; therapy.
KW	

```

XX New isolated protein, Zins1 - obtained from human placenta, which
PT increases the proliferation of pancreatic islet cells, used for
PT treating diabetes
XX
XX Example 1; Page 62-63; 77pp; English.
XX
XX This sequence is a N-terminally tagged version of the human Zins1
CC protein of the invention. The Zins1 protein was isolated from human
CC placenta, and is believed to be a new version of the mature protein of
CC placenta, having disulphide bonded A and B chains. The protein can be
CC used for stimulating the proliferation of pancreatic islets to increase
CC insulin secretory capacity of mammals. In particular it can be used for
CC the treatment of diabetes. It can also be used for stimulating in vitro
CC proliferation of pancreatic islet cells. It can also be used for
CC production of antibodies and in detection and diagnosis.
XX
XX Sequence 159 AA;
SQ
Alignment Scores:
Pred. No.: 4,59e-66 Length: 159
Score: 619.00 Matches: 114
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 79.97% Indels: 0
DB: 19 Gaps: 0
US-09-518-842-1 (1-420) x AAN69170 (1-159)
QY 76 GCAGAGCTGAGGGATGCTGCTCCGATTTGGAAACACTTCTGTCATATTGCCCATG 135
Db 46 AlaGluLeuArgGlyCysGlyProArgPheGlyLysHisLeuLeuSerTyrCysPromet 65
2Y 136 CQTGAGAGACATTCACACACCCAGAGGGTGGCTGCTGGAACTGAGCGTCCCAAA 195
Db 66 ProGluLysThrPheThrThrProGlyGlyTrpLeuLeuGluSerGlyArgProlys 85
2Y 196 GAATGGTGTCACTCCACACACAAAGATGGACAGCCTTAGGTACGACATCAGAATTC 255
Db 86 GluMetValSerThrSerAsnAsnLysAspGlyGlnAlaLeuGlyThrThrSerGluPhe 105
2Y 256 ATTCCTAATTGTGACAGAGTGAAGAACACCTGTCTGAAAGGCGACGCATCATTTGAAG 315
Db 106 IleProAsnLeuSerProGluLeuLysLysProLeuSerGluGlyGlnProSerLeuLys 125
2Y 316 ABAATAATCTTCCCGCAAAAGACAGTGGACGCTCAGATTTGATCCATTCTGTGTGT 375
Db 126 LysIleLeuSerArgLysLysArgSerGlyArgHisArgPheAspProPheCysCys 145
2Y 376 GAAGTAATTGTGACGATGGAATTCAGTTAAATATTATGTACA 417
Db 146 GluValIleCysAspGlyThrSerValLysLeuCysThr 159
RESULT 8
D AAP30392 standard; Protein; 182 AA.
X
X AAP30392;
X
X 25-MAR-2003 (updated)
X 25-MAY-1992 (first entry)
X
X Sequence of porcine preprorelaxin.
X Relaxin; hormone.
X
X Sus scrofa domestica.
X EP86649-A.
X
X 24-AUG-1983.
X
X 11-FEB-1983; 83EP-0300714.
X

```

```

XX
PR 12-FEB-1982; 82AU-0002695.
PR 11-FEB-1983; 83AU-0011834.
XX
XX (FLOR-) FLOREY INST EXP PHY.
PA (FLOR-) FLOREY HOWARD INST.
XX
XX Hudson PJ, Haley JD, Niall HD, Shine J;
XX WPI: 1983-748587/35.
DR N-PSDB; AAN30196.
XX
XX Genes and DNA transfer vectors for prorelaxin expression - useful
PT in prodd. of porcine relaxin for veterinary and human use
XX
XX Disclosure; Fig 5; 50pp; English.
XX
XX The inventors claim synthetic porcine preprorelaxin and prorelaxin
CC and synthetic A, B and C peptide chains of prorelaxin, and a gene for
CC expression of porcine preprorelaxin or prorelaxin, and their sub-
CC units (see AAN30196). They also claim a double-stranded DNA fragment
CC for the expression of the signal peptide chain of porcine
CC preprorelaxin comprising a coding strand and a complementary strand
CC corresp. to a defined mRNA sequence (see AAN30187-N30194) which
CC corresp. to the most homologous regions between the pig and rat cDNA
CC sequences. A probe (AAN30195) is also claimed.
CC (Updated on 25-MAR-2003 to correct PF field.)
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 182 AA;
Alignment Scores:
Pred. No.: 8,98e-13 Length: 182
Score: 180.00 Matches: 61
Percent Similarity: 38.02% Conservative: 12
Best Local Similarity: 31.77% Mismatches: 55
Query Match: 23.26% Indels: 64
DB: 4 Gaps: 6
US-09-518-842-1 (1-420) x AAP30392 (1-182)
QY 1 ATGGCCAGCTGTCGGTCTCTATCTGCCAGCAATCTGGCTGCTGCTGAGCAACTCTCT 60
Db 1 MetProArgLeuPhe---SerTyrLeuLeuGlyValTrpLeuLeuLeuSerGlnLeuPro 19
QY 61 AGAGAA-----AGCCTAGCAGCAGCTGAGGGATGTGGCCCGATTTCGA 108
Db 20 ArgGluLeuProGlyGlnSerThrAsnAspPheIleLysAlaCysGlyArgGluLeuVal 39
QY 109 AACACCTTGTCTCATATTGCCCCCATGCTGAGAGACATTCACCCACCCAGAGGG 168
Db 40 ArgLeuTrpValGluIleCysGlySerVal-----Ser 50
QY 169 TGG-----CTGCTGGAATCTGGAGCTGCCAAGAA 198
Db 51 TrpGlyArgThrAlaLeuSerLeuGluProGlnLeuGluThrGlyProProAlaGlu 70
QY 199 ATGGTGTCAACCTCCCAACCAAGATGGACAGCTTAGGTACGACATCAGATTCATT 258
Db 71 ThrMetProSerSerIleThrLysAspAlaGluIleLeuLysMetMetLeuGluPheVal 90
QY 259 CCTAATTGTCTACAGAGCTGAAGAAACCATCTGTCTGAAGGCGACCATTCATTG 312
Db 91 ProAsnLeuProGlnGluLeuLysAlaThrLeuSerGluArgGlnProSerLeuArgGlu 110
QY 313 -----AAGAAATATA 324
Db 111 LeuGlnGlnSerAlaSerLysAspSerAsnLeuAsnPheGluGluPheLysIleIle 130
QY 325 CTT----- 327
Db 131 LeuAsnArgGlnAsnGluAlaGluAspLysSerLeuLeuLeuLysAsnLeuGlyLeu 150

```

328 -----TCCGCAAAAGAGAGTGGAGCTCACAGATTTCATTCATTCGTGTGAA 378
 151 AspLysHisSerArgLysArgLeuPheArgMetThrLeuSerGluLysCysGln 170
 379 GTAATTTCGTGACGATGGAACCTTCAGTTAAATTATGT 414
 171 ValGlyCysIleArgLysAspIleAlaArgLeuCys 182

RESULT 9

AAP40156
 D AAP40156 standard; Protein; 182 AA.

AAP40156;

25-MAR-2003 (updated)
 11-FEB-1992 (first entry)

Sequence of porcine preprorelaxin.

Labour; birth; hormone; relaxin.

Fig.

Key Location/Qualifiers
 Peptide 1..25
 /label= signal
 Peptide 26..57
 /label= B-chain
 Peptide 58..160
 /label= C-peptide
 Peptide 161..182
 /label= A-chain

-EP101309-A.

22-FEB-1984.

11-AUG-1983; 83EP-0304662.

12-AUG-1982; 82AU-0005352.

11-AUG-1983; 83AU-0021906.

01-JAN-1988; 88EP-0104503.

(FLOR-) FLOREY HOWARD INST.

Hudson PJ, Shine J, Niall HD, Tregear GW;

WPI; 1984-050918/09.

N-PSDB; AAN40125, AAN40126.

Genes for human relaxin, preprorelaxin and preprorelaxin prodn. -

prepd. by recombinant DNA techniques

Disclosure; Fig 3; S1pp; English.

The inventors claim the gene for the expression of human
 preprorelaxin (HPP) and its sub-units. Also claimed are synthetic
 HPP, HP, fragments and analogues. Human relaxin and its analogues
 are prepd. for therapeutic purposes, esp. in clinical intervention
 in cases of difficult labour.

(Updated on 25-MAR-2003 to correct PA field.)

Sequence 182 AA;

Lignment Scores:
 red. No.: 6.35e-12 Length: 182
 core: 173.00 Matches: 60
 percent Similarity: 37.50% Conservative: 12
 est Local Similarity: 31.25% Mismatches: 56
 uery Match: 22.35% Indels: 64
 B: 5 Gaps: 6

S-09-518-842-1 (1-420) x AAP40156 (1-182)

QY 1 ATGGCCAGCCGTTCGCGTCTCTATCTGCCACAATCTGGCTGCTGAGCAACTCCTT 60
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 1 MetProArgLeuPhe---Ser***LeuLeuGlyValTrpLeuLeuLeuSerGlnLeuPro 19
 QY 61 AGAGAA-----AGCCTAGCACAGAGCTGAGGGATGTGGTCCCGATTTGGA 108
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 20 ArgGluIleProGlyGlnSerThrAsnAspPheIleLysAlaCysGlyArgGluLeuVal 39
 QY 109 AAACACTTGTCTCATATTGCCCTCCCTGCTGAGAAGACATTCACCACCCAGGAGG 168
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 40 ArgLeuTrpValGluIleCysGlySerVal-----Ser 50
 QY 169 TGG-----CTGCTGGAATCTGGAGCTCCCAAGAA 198
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 51 TrpGlyArgThrAlaLeuSerLeuGluProGlnLeuGluThrGlyProProAlaGlu 70
 QY 199 ATGGTGTCAACCTCCACACAAAGATGGAACGCTTAGGTACGACATCAGAATTCAT 258
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 71 ThrMetProSerSerIleThrLysAspAlaGluIleLeuLysMetMetLeuGluPheVal 90
 QY 259 CCTAATTTGTACACAGCTGAGAAACCACTGTCTGAAGGCGAGCCATCATTTG----- 312
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 91 ProAsnLeuProGlnGluLeuLysAlaThrLeuSerGluArgGlnProSerLeuArgGlu 110
 QY 313 -----AAGAAATAATA 324
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 111 LeuGlnGlnSerAlaSerLysAspSerAsnLeuAsnPheGluGluPheLysLysIlelle 130
 QY 325 CTT----- 327
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 131 LeuAsnArgGlnAsnGluAlaGluAspLysSerLeuLeuGluLeuLysAsnLeuGlyLeu 150
 QY 328 -----TCCGCAAAAGAGAGTGGAGCTCACAGATTTCATTCGTGTGAA 378
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 151 AspLysHisSerArgLysArgLeuPheArgMetThrLeuSerGluLysCysGln 170
 QY 379 GTAATTTCGTGACGATGGAACCTTCAGTTAAATTATGT 414
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 171 ValGlyCysIleArgLysAspIleAlaArgLeuCys 182

RESULT 10

AAP94621

ID AAP94621 standard; protein; 185 AA.

XX AC AAP94621;

XX XX

DT 25-MAR-2003 (updated)

DT 21-JUN-1990 (first entry)

DE Amino acid sequence of human preprorelaxin H2.

XX DE Relaxin; H2-relaxin; pubic symphysis.

XX KM Relaxin; H2-relaxin; pubic symphysis.

XX OS Homo sapiens.

XX XX Key Location/Qualifiers

FX Peptide 1..25 /label=Signal peptide

FT Peptide 26..57 /label=B-chain

FT Peptide 58..161 /label=C-peptide

FT Peptide 162..159 /label=A-chain

FT FT

XX EP303033-A.

PN 15-FEB-1989.

PD 12-DEC-1983; 88EP-0110103.

XX 13-DEC-1982; 82AU-0007247.

PR

XX (FLOR-) FLOREY INST EXPERIMENTAL PHYSIOLOGY.
 PA Hudson PJ, Tregear GW, Niall HD;
 XX WPI: 1989-047874/07.
 XX N-PSDB; AAN92483.
 XX
 PT New human H2-relaxin analogues -
 PT with shortened and/or modified A and/or B chains.
 XX Disclosure; Page -: 25pp; English.
 XX
 XX "H2" relaxin is deduced from a cDNA clone, has the general properties of
 CC a growth factor and is capable of altering nature of connective tissue
 CC and inducing smooth muscle contraction, specifically during labour.
 CC (Updated on 25-MAR-2003 to correct PF field.)
 CC (Updated on 25-MAR-2003 to correct PA field.)
 CC (Updated on 25-MAR-2003 to correct PI field.)
 XX SQ Sequence 185 AA;

Alignment Scores:
 Pred. No.: 6.48e-06 Length: 185
 Score: 123.50 Matches: 53
 Percent Similarity: 35.35% Conservative: 17
 Best Local Similarity: 26.77% Mismatches: 55
 Query Match: 15.96% Indels: 73
 DB: 10 Gaps: 7

JS-09-518-842-1 (1-420) x AAP94621 (1-185)
 Y 1 ATGGCCAGCCGTTCGGTCTCTATCGCAGCATCTGGCTGCTGCTGAGCAACTCCTT 60
 b 1 MetProArgLeuPhePheHisLeuLeuGlyValCysLeuLeuLeuAsnGlnPheSer 20
 Y 61 -----AGAGAAAGCCTA 72
 b 21 ArgAlaValAlaAspSerTirpMetGluGluValIleLeuLeuLeuGlyArgGluLeuVal 40
 Y 73 GCAGCAGAGCTGAGGGATGTGGT---CCCGATTGGAAACACCTGCTGCTATATGC 129
 b 41 ArgAlaGlnIleAlaIleCysGlyMetSerThrTrpSerLysArgSerLeuSerGlnGlu 60
 Y 130 CCCATGCTGAGACACATTCACCACCCAGGAGGCTGCTGGAATCTGGACGT 189
 b 61 AspAlaProGln-----ThrPro-----Arg 67
 Y 190 CCC---AAGAAATGGTGTCACTCCCAACAAAGATGACAGCCTTAGGTACGACA 246
 b 68 ProValAlaGluIleValProSerPheIleAsnLysAspThrGluThrIleAsnMetMet 87
 Y 247 TCAGAAATTCATTCCTAATTTGTCACCCAGCTGAGAAACCACTGCTGAGGCGACCA 306
 D 88 SerGluPheValAlaAsnMetProGlnGluLeuLysLeuThrLeuSerGluMetGlnPro 107
 Y 307 TCATTG----- 312
 b 108 AlaLeuProGlnLeuGlnHisValProValLeuLysAspSerSerLeuLeuPheGlu 127
 Y 313 -----AGAAATATATCTT----- 327
 b 128 GluPheLysLysLeuIleArgAsnArgGlnSerGluAlaAlaAspSerProSerGlu 147
 Y 328 -----TCCCGCAAAAGAGAGAGTGGACGTCACAGATT 360
 b 148 LeuLysTyrLeuGlyLeuAspThrHisSerArgLysLysArgGlnLeuTyrSerAlaLeu 167
 Y 361 GATCCATCTGTGTGAAGTAATTTGTGACCATGCACTTCAGTTAAATATATG 414
 b 168 AlaAsnLysCysCysHisValGlyCysThrLysArgSerLeuAlaArgPheCys 185

ESULT 11

AAP40108
 ID AAP40108 standard; Protein; 185 AA.
 XX
 AC AAP40108;
 XX
 DT 25-MAR-2003 (updated)
 DT 04-FEB-1992 (first entry)
 XX
 DE Sequence of human preprorelaxin H2.
 XX Relaxin; hormone; probe; uterine contraction.
 KW
 OS H2.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..24 /label= signal
 FT Region 25..57 /label= B-chain
 FT Region 58..161 /label= C-peptide
 FT Region 162..185 /label= A-chain
 FT
 XX EF112149-A.
 XX
 PD 27-JUN-1984.
 XX
 PF 12-DEC-1983; 83EP-0307553.
 XX
 PR 13-DEC-1982; 82AU-0007247.
 PR 09-DEC-1983; 83AU-0022283.
 PR 01-JAN-1988; 88EP-0110103.
 XX
 PA (FLOR-) FLOREY INST EXPERIMENTAL PHYSIOLOGY.
 XX
 PI Hudson PJ, Niall HD, Tregear GW;
 DR WPI: 1984-160129/26.
 DR N-PSDB; AAN40086.
 XX
 PT Gene for expression of human H2-preprorelaxin - useful as vector
 PT for prodn. of the peptide after expression etc.
 XX
 PS Disclosure; Fig 2; 56pp; English.
 XX
 CC The inventors claim a ds DNA SQ which comprises a coding strand and
 CC a complementary strand corresp. to (AAN40086) and its subunits. Also
 CC claimed are synthetic human H2-preprorelaxin, human H2-prorelaxin
 CC and human H2-relaxin together with their analogues having shortened A
 CC or B chains, or modified to contain different AA residues.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX
 SQ Sequence 185 AA;

Alignment Scores:
 Pred. No.: 1.5e-05 Length: 185
 Score: 120.50 Matches: 49
 Percent Similarity: 38.38% Conservative: 22
 Best Local Similarity: 26.44% Mismatches: 67
 Query Match: 15.57% Indels: 47
 DB: 5 Gaps: 5

US-09-518-842-1 (1-420) x AAP40108 (1-185)
 QY 1 ATGGCCAGCTGTTCCGGTCTCTATCGCAGCAATCTGGCTGCTGAGCAACTCCTT 60
 b 1 MetProArgLeuPhePheHisLeuLeuGlyValCysLeuLeuLeuAsnGlnPheSer 20
 QY 61 AGA-----GAAAGCTAGCAGCAGAG---CTGAGGGGATGTGGTCCCGATTGGA 108
 b 21 ArgAlaValAlaAspSerTirpMetGluGluValIleLysLeuCysGlyArgGluLeuVal 40

109 AAACACTTGCTGTCATATTGCCCATGCTGCTGAGAGACATTCACACCCAGGAGG 168
 41 ArgAlaGlnIleAlaIleCysGlyMetSerThrTrpSerLysArgSerLeuSerGlnGlu 60
 169 TGGCTGCTGGAATCTGGACGCTCC--AAAGAAATGGTGTCAACCTCCCAACAAAGAT 225
 61 AspAlaProGlnAsnProArgProValAlaGluIleValProSerPheIleAsnLysAsp 80
 226 GGCACACCTTAGTACGACATCAGATTCCTTCAATTCCTTCAATTCCTTCAATTCCTT 285
 81 ThrGluThrIleAsnMetMetSerGluPheValAlaAsnLeuProGlnGluLeuLysLeu 100
 286 CCACTGTCTCAAGCGGACCATCATG----- 312
 101 ThrLeuSerGluMetGlnProAlaLeuProGlnLeuGlnGlnHisValProValLeuLys 120
 313 -----AAGAAATAATACTT----- 327
 121 AspSerSerLeuLeuPheGluGluPheLysLysLeuIleArgAsnArgGlnSerGluAla 140
 328 -----TCCCGCAAAAG 339
 141 AlaAspSerSerProSerGluLeuLysTyrLeuGlyLeuAspThrHisSerArgLysLys 160
 340 AGAAGTGGAGCTCACAGATTGATCCATTCTGTGTGAAGTAAATTGTGACCATGGAAT 399
 161 ArgGlnLeuTyrSerAlaLeuAlaAsnLysCysHisValGlyCysThrLysArgSer 180
 400 TCAGTTAAATTATGT 414
 181 LeuAlaArgPheCys 185

ESULT 12

AP40154
 D AAP40154 standard; Protein; 185 AA.

K AAP40154;

K 25-MAR-2003 (updated)

T 11-FEB-1992 (first entry)

T Sequence of human preprorelaxin.

K Labour; birth; hormone; relaxin.

K Homo sapiens.

K Key Location/Qualifiers

T Peptide 1..25

T /label= signal

T Peptide 26..57

T /label= B-chain

T Peptide 58..161

T /label= C-peptide

T Peptide 162..185

T /label= A-chain

K EP101309-A.

K 22-FEB-1994.

K 11-AUG-1983; 83EP-0304662.

K 12-AUG-1982; 82AU-0005352.

K 11-AUG-1983; 83AU-0017906.

K 01-JAN-1988; 88EP-0104503.

K (FLOR-) FLOREY HOWARD INST.

K Hudson PJ, Shine J, Niall HD, Tregear GW;

K WPI; 1984-050918/09.

K N-PSDB; AAP40101.

XX Genes for human relaxin, prorelaxin and preprorelaxin prodn. -
 PT prepd. by recombinant DNA techniques
 XX Disclosure; Fig 2; 51pp; English.
 CC The inventors claim the gene for the expression of human
 CC preprorelaxin (HPP) and its sub-units. Also claimed are synthetic
 CC HPP, HP, fragments and analogues. Human relaxin and its analogues
 CC are prepd. for therapeutic purposes, esp. in clinical intervention
 CC in cases of difficult labour.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX Sequence 185 AA;

Alignment Scores:

Pred. No.: 1.99e-05 Length: 185
 Score: 119.50 Matches: 52
 Percent Similarity: 37.11% Conservative: 20
 Best Local Similarity: 26.80% Mismatches: 57
 Query Match: 15.44% Indels: 65
 DB: 5 Gaps: 7

US-09-518-842-1 (1-420) x AAP40154 (1-185)

QY 1 ATGCCAGCCTGTTCCGGTCTCTATCTGCCAGCAATCTGCTGCTGCTGAGCCAACTCCTT 60
 DB 1 MetProArgLeuPheLeuPheHisLeuLeuGluPheCysLeuLeuLeuAsnGlnPheSer 20
 QY 61 AGAGAAAGCCTAGCAGCAGAG-----CTGAGGGGATGTGTCCTCCCATTT 105
 DB 21 Arg--AlaValAlaAlaLysTrpLysAspValIleLysLeuCysGlyArgGluLeu 39
 QY 106 GCAAAACACTGCTGCTCATATTGCCCCATG-----CCTGAG 141
 DB 40 ValArgAlaGlnIleAlaIleCysGlyMetSerThrTrpSerLysArgSerLeuSerGln 59
 QY 142 AAGACATTCCACCACCCAGGAGGGTGGTCTGGAATCTGGAGCTGCC---AAGAA 198
 DB 60 GluAspAlaProGlnThrPro-----ArgProValAlaGlu 71
 QY 199 ATGGTGCAACCTCCAAACAACAAGATGGACAGCCTTAGTACGACATCAGATTCATT 258
 DB 72 IleValProSerPheIleAsnLysAspThrGluThrIleIleMetLeuGluPheIle 91
 QY 259 CCTAATTGTCCACGAGCTCAAGAACCACTCTCTGAGGGGAGCCCATCTG----- 312
 DB 92 AlaAsnLeuProProGluLeuLysAlaLeuSerGluArgGlnProSerLeuProGlu 111
 QY 313 -----AAGAAA 318
 DB 112 LeuGlnGlnTyrValProAlaLeuLysAspSerAsnLeuSerPheGluGluPheLys 131
 QY 319 ATAATCTTCCCGC----- 333
 DB 132 LeuIleArgAsnArgGlnSerGluAlaAlaAspSerAsnProSerGluLeuLysTyrLeu 151
 QY 334 -----AAAAAGGAAGTGGACGTCACAGATTTGATCCATCTGT 372
 DB 152 GlyLeuAspThrHisSerGlnLysLysArgArgProTyrValAlaLeuPheGluLysCys 171
 QY 373 TGTGAAGTAAATTGTGACGATGGAACTTCAGTTAAATATGT 414
 DB 172 CysLeuIleGlyCysThrLysArgSerLeuAlaLysTyrCys 185

RESULT 13

AAP40155

ID AAP40155 standard; Protein; 185 AA.

XX AAP40155;

XX 25-MAR-2003 (updated)

DT 11-FEB-1992 (first entry)

prostate cancer-associated polynucleotides (designated PC genes) that selectively hybridise to a sequence that is at least 80% identical to them. The prostate cancer-associated polynucleotide sequences are differentially expressed in prostate tumour tissue or in prostate cancer and are derived from the tissues of various organisms such as humans or other mammals (e.g. mice, sheep and dogs). The methods of the invention are useful for diagnosing and treating prostate cancer in mammals. The prostate cancer-associated genes are useful for diagnosing or treating prostate cancer, as well as for identifying modulators of prostate cancer or agents that inhibit prostate cancer. The nucleic acid sequences are particularly useful in gene therapy, as a vaccine or in antisense applications.

ABG61800-ABG61944 represent prostate cancer-associated proteins.

Sequence 185 AA;

Alignment Scores:
 red. No.: 1.98e-05 Length: 185
 Score: 119.50 Matches: 52
 Percent Similarity: 37.11% Conservative: 20
 Best Local Similarity: 26.80% Mismatches: 57
 Query Match: 15.44% Indels: 65
 Gaps: 7

3-09-518-842-1 (1-420) x ABG61819 (1-185)

1 ATGGCCAGCTGTTCGGTCCCTATCTGCCAGCAATCTGCTGCTGAGCAACTCCTT 60
 1 MetProArgLeuPheLeuPheHisLeuLeuGluPheCysLeuLeuLeuAsnGlnPheSer 20
 61 AGAAGAGCTAGCAGCAGAG-----CTGAGGGGATGTGTCCTCCCGATTT 105
 21 Arg--AlaValAlaAlaLysTrpLysAspValIleLeuLeuCysGlyArgGluLeu 39
 106 GGAACACTTGTCTCATATTCCTCCCATG-----CCTGAG 141
 40 ValArgAlaGlnIleAlaIleCysGlyMetSerThrTrpSerLysArgSerLeuSerGln 59
 142 AACACATCCACCACCCAGGAGGGTGGCTGCTGGAATCTGGAGCTGCC---AAGAA 198
 60 GluAspAlaProGlnThrPro-----ArgProValAlaGlu 71
 199 ATGGTGTCACTCCCAACAACAAGATGGACAGCCTTAGTGACGACATCAGAATTCATT 258
 72 IleValProSerPheIleAsnLysAspThrGluThrIleIleIleMetLeuGluPheIle 91
 259 CCTAATTTGTCACAGCTGAGAACCACTGTCTGAGGGGAGCCATCTG----- 312
 92 AlaAsnLeuProGluLeuLysAlaAlaLeuSerGluArgGlnProSerLeuProGlu 111
 313 -----AAGAAA 318
 112 LeuGlnGlnTyrValProAlaLeuLysAspSerAsnLeuSerPheGluGluPheLysLys 131
 319 ATAATCTTCCGCG----- 333
 132 LeuIleArgAsnArgGlnSerGluAlaAlaAspSerAsnProSerGluLeuLysTyrLeu 151
 334 -----AAGAGAGAGTGGAGCTGACAGATTCATTCATTCCT 372
 152 GlyLeuAspThrHisSerGlnLysLysArgArgProTyrValAlaLeuPheGluLysCys 171
 373 TGTCAGTAATTTGTGACGATGGAACCTTCAGTTAAATATGT 414
 172 CysLeuIleGlyCysThrLysArgSerLeuAlaLysTyrCys 185

SULT 15
 UR64904
 AAR64904 standard; Protein; 150 AA.
 AAR64904;
 25-MAR-2003 (updated)

06-SEP-1995 (first entry)
 XX Prorelaxin fragment amino acids 12-161.
 DE Prorelaxin; cervical ripening; ovarian peptide hormone;
 XX mammary gland development; skin elasticity; cardiovascular therapy;
 KW relaxin; premature labour; plasmid pTR21.
 KW Synthetic.
 OS W09500645-A2.
 XX W09500645-A2.
 XX 05-JAN-1995.
 XX 20-JUN-1994; 94WO-US06997.
 XX 21-JUN-1993; 93US-0080354.
 XX (GETH) GENENTECH INC.
 PI Breese T, Hayenga K, Rinderknecht E, Vandlen R;
 PI Yansura D;
 XX WPI; 1995-052082/07.
 DR N-PSDB; AAQ76309.
 XX Relaxin prodn. from non-natural recombinant prorelaxin - by
 PT cleavage of leader and C peptide chain, also new prorelaxin and
 PT related DNA, vectors etc.
 XX Example 1; Fig 9A; 62pp; English.
 PS AAQ76309 is the Not I-Bam HI restriction fragment of the plasmid
 CC pTR21, which encodes AAR64904, prorelaxin amino acids 12-161.
 CC Relaxin is produced by removing the non-naturally occurring leader
 CC and C-peptide from PR with a cleavage agent. Relaxin is an ovarian
 CC peptide hormone involved in the inhibition of premature labour,
 CC cervical ripening and the development of the mammary glands. It
 CC may also improve skin elasticity and has been used in cardiovascular
 CC therapy.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX SQ Sequence 150 AA;

Alignment Scores:
 Pred. No.: 9.74e-05 Length: 150
 Score: 113.50 Matches: 45
 Percent Similarity: 36.42% Conservative: 14
 Best Local Similarity: 27.78% Mismatches: 46
 Query Match: 14.66% Indels: 57
 DB: 16 Gaps: 6
 US-09-518-842-1 (1-420) x AAR64904 (1-150)
 QY 61 AGAAGAGCTAGCAGCAGCAGCTGAGGGATGTGGT---CCCCGATTTGAAACACTTG 117
 DB 2 ArgGluLeuValArgAlaGlnIleAlaIleCysGlyMetSerThrTrpSerLysArgSer 21
 QY 118 CTGTCTATTGCCCCATGCTGCTGAGAGACATTCACCACCCAGGAGGGTGGCTGCTG 177
 DB 22 LeuSerGlnGluAspAlaProGln-----ThrPro----- 31
 QY 178 GAATCTGGAGCTCCC---AAGAAATGGTGTCAACCTCCACAAACAAAGATGCAAGCC 234
 DB 32 -----ArgProValAlaGluIleValProSerPheIleAsnLysAspThrGluThr 48
 QY 235 TTAGGTACGACATCAGAAATTCATTCTTAATTTGTACCAGAGCTGAAGAACACCTGTCT 294
 DB 49 IleAsnMetMetSerGluPheValAlaAsnLeuProGlnGluLeuLysLeuThrLeuSer 68
 QY 295 GAAGGGAGCCCATCATG----- 312
 DB 69 GluMetGlnProAlaLeuProGlnLeuGlnGlnHisValProValLeuLysAspSerSer 88

```

313 -----AAGAAAATAATACTT----- 327
b 89 LeuLeuPheGluGluPheLysLeuIleArgAsnArgGlnSerGluAlaAlaAspSer 108
y 328 -----TCCCGCAAAAGAGAGACTGGA 348
b 109 SerProSerGluLeuLysTyrLeuGlyLeuAspThrHisSerArgLysLysArgGlnLeu 128
y 349 CGTCACAGATTGATCCATTCTGTTGTGAAGTAATTGTGACGACGAACTTCAGTTAAA 408
b 129 TyrSerAlaLeuAlaAsnLysCysHisValGlyCysThrLysArgSerLeuAlaArg 148
y 409 TTATGT 414
b 149 PheCys 150

```

search completed: October 9, 2003, 12:15:11
 ob time : 65.5 secs